Cleft sentences, construction grammar and grammaticalization

Amanda L. Patten

Thesis submitted for the degree of Doctor of Philosophy
Linguistics and English Language
The University of Edinburgh
2010

Abstract

This thesis examines the structure and function of the English *it*-cleft configuration within the framework of construction grammar. My analysis begins with the claim that *it*-clefts are a subtype of specificational copular sentence. After identifying problems with previous accounts, I outline my own, original analysis of specificational *NP be NP* sentences. I argue that specificational meaning involves an asymmetric predication relation and is dependent upon the inherent semantics of definite noun phrases (rather than syntactic movement). I treat nominal predication set theoretically, as a semantic relation between members and sets. I claim that specificational meaning is brought about by a reinterpretation of the class-membership relation involving definite NP predicates, whereby the referent is identified as the unique member of a restricted and existentially presupposed set.

As a member of the family of specificational copular sentences, the *it*-cleft inherits properties from the more basic construction. From this, it follows that *it*-clefts should also involve a nominal predication relation, containing a definite NP predicate. This leads me to argue in favour of a non-derivational extraposition-from-NP analysis of *it*-clefts, in which the pronoun *it* and the cleft clause (analysed here as a restrictive relative) function together as a discontinuous definite description. My analysis improves on similar accounts of this type in two ways. First, since my analysis explains the role that definite descriptions play in the creation of specificational meaning, I am able to explain, rather than simply identify, the numerous similarities between *it*-clefts and definite noun phrases. Second, my analysis of specificational sentences as involving a nominal predication relation allows for a straightforward account of the relationship between specificational and predicational *it*-clefts.

The thesis also examines the historical development of the *it*-cleft construction. I show that (a) much of the *it*-cleft's structure is reminiscent of an earlier stage of the language and (b) the construction has become increasingly schematic and productive over time, sanctioning instances which override inheritance from the more basic specificational schema. In this way, the historical evidence provides an explanation for the *it*-cleft's idiosyncratic properties. Together, my synchronic and diachronic analyses add up to a maximally explanatory account of the *it*-cleft construction.

Declaration

I declare that this thesis has been composed by me and that the work contained within is my own, except where explicitly stated in the text. None of this material has previously been submitted for another degree or professional qualification.

Amanda Patten

Acknowledgements

I would first like to thank Nikolas Gisborne, my PhD supervisor, for his constant encouragement and for his time. His helpful comments and suggestions have forced me to work through my ideas and have continually guided me into new and interesting areas of research. I am immensely grateful to Nik for getting me started on the topic of clefts and for helping me to formulate a workable proposal.

I would also like to thank Claire Cowie and April McMahon for encouraging me to make the move to Edinburgh in the first place. I have profited from the active research community at Edinburgh and I would like to thank the members of the English Language Research Group, the Syntax and Semantics Research Group and the Construction Grammar Reading Group for sharing their ideas and their research. In particular, I would like to thank Caroline Heycock, Graeme Trousdale and Dan Wedgwood for discussing my work with me and also Linda van Bergen for her expertise in interpreting the historical data.

Thanks also to Katie Keltie and Kirsty Woomble, members of the administrative staff at the University of Edinburgh, who were always on hand to answer my questions. My fellow postgraduates also deserve thanks for their support and friendship – Rhona Alcorn, Sarah Anderson (nee Collie), Will Barras, Remco Knooihuizen, Marleen Spaargaren and especially Lynn Clark. As my officemate and friend, Lynn has been with me from the beginning – thanks for putting up with me.

Special thanks go to Elizabeth Traugott and Graeme Trousdale for showing interest in my work and for inviting me to participate in the *Gradience, gradualness and grammaticalization* workshop at the NRG4 conference in Leuven. I am grateful to Elizabeth and Graeme and to my fellow contributors to the *GGG* volume for their useful comments on my written up paper. I would also like to thank the audiences at the ICLCE2 in Toulouse, the LAGB meeting in London 2007, the NRG4 in Leuven and Cleft08 in Berlin.

I would like to thank Roger Higgins for taking the time to talk about clefts with me early on in my studies and for sharing his enthusiasm for the topic with me. Thanks also to Cathy Ball, Marcel Den Dikken and Nancy Hedberg for helping me to get access to unpublished or not widely distributed works. I must also thank Susan Pintzuk, Ann Taylor and Anthony Warner for introducing me to the Penn-Helsinki Parsed Corpora and for guidance in using the CorpusSearch 2 search engine.

I would especially like to thank the AHRC for funding this project. Without this financial support, I would not have been able to continue in education. Finally, I would like to thank my parents for their never-ending support (financial and otherwise) and for caring much more about me than about this PhD.

Contents

1. INTRODUCTION AND BACKGROUND

1.1 An outline of the project

This thesis examines the synchronic structure and diachronic development of the English *it*-cleft within the framework of construction grammar. *It*-clefts are interesting for a number of reasons. For one thing, they have a non-standard structure which appears not to conform to the general rules of the language. If we take a look at the example in (1), we can see that *it*-clefts have four main components: the introductory pronominal *it*, a form of the copular verb *be*, a postcopular phrasal element and a sentence-final clause.

(1) [It] [was] [Howard] [that left]

From this example, we can see that the *it*-cleft's syntactic configuration is difficult to make sense of. The *that*-clause is structured internally like a restrictive relative. However, proper names, such as *Howard*, are full noun phrases. As such, they cannot normally be modified by restrictive relative clauses. So how does this clause relate to the postcopular element, if at all? Can we really call this a restrictive relative clause? If so, what does it modify? If not, are we dealing with a clausal structure that is unique to the *it*-cleft? Equally problematic is the role of initial *it*. Is this an expletive dummy subject and if so, why is it there? Does it operate as a syntactic placemarker and if so, for which element? Or is the constituent *it* related in a different way to other elements in the sentence?

In addition, *it*-clefts have a number of unusual semantic, pragmatic and discourse-functional properties. These are particularly interesting since it is not immediately clear which elements in the cleft structure contribute to the meaning of the construction. For example, the *it*-cleft is a focusing construction. The primary informational content is placed in the syntactically marked postcopular focal position and is often given primary stress (see (2)). However, it is not at all obvious why this

particular syntactic configuration should be chosen as a focusing device. Is this its primary function?[1]

(2) It was HOWARD that left

A further property of the *it*-cleft is that it exhibits an exhaustiveness implicature. For example, in (2) we assume that *Howard* was the only person that left on that occasion. They are also presuppositional; the information in the sentence-final clause is not asserted and is preserved under negation. For instance, in example (3), we are told that *Howard didn't leave* but we are left with the presupposition that *somebody (else) did leave*. This begs the question, where do these pragmatic meanings come from? Which elements contribute to them?

(3) It wasn't Howard that left

Cleft sentences also have a specificational (or identifying) meaning. For some authors specificational meaning involves a 'value-variable' relationship (see especially Declerck 1988). For example, (2) identifies *Howard* as the value for the variable *x* in the proposition expressed by the sentence-final clause, *x left*. However, for others, specificational meaning is attributed to a special use of the copular verb. So does *be* have a specificational meaning in the *it*-cleft? If not, where does the specificational meaning of *it*-clefts come from?

To a large extent, how these questions are answered (and perhaps whether they are even asked at all) depends upon how we think *it*-clefts relate to other constructions, or configurations, in the language. Most approaches to *it*-clefts fall into two broad categories: those that understand *it*-clefts in relation to simple subject-predicate

[1] I use the term 'focus' to refer to a unit of information structure where the assertion differs from the presupposition (see Lambrecht 1994; Vallduví and Engdahl 1996; Erteschik-Shir 2007). Although focus is often marked by intonation, Lambrecht (1994: 208) observes that "accent placement and focus marking are not to be equated". Where focus marking is unclear or ambiguous in my examples, I make use of small capitals to indicate the marking of focus by intonation.

sentences, such as (4), and those that relate *it*-clefts to other specificational copular sentences, such as (5).

(4) Howard left

(5) The one that left was Howard

I discuss these two approaches and the analyses that result from them in §1.2. Essentially, justification for the first approach comes from the truth-conditional equivalence between *it*-clefts and simple subject-predicate sentences. From this perspective, *it*-clefts are viewed primarily as a means of marking focus syntactically. The second approach, on the other hand, builds on the fact that the *it*-cleft is a copular construction with a specificational meaning. So which is the right approach? The answer to this question depends upon what we think is the primary function of *it*-clefts and asking which perspective can best explain the range of properties that *it*-clefts display.

There are also different varieties of *it*-cleft which are sometimes regarded as separate structures. One domain of variability involves what can occur in the focus position. Although, most frequently, the focal element is a noun phrase, *it*-clefts permit a range of elements as the complement of *be*, such as the prepositional phrase in (6). Should these examples be analysed in the same way as those with nominal foci or do they require a separate analysis? Can the sentence-final clause still be analysed as a restrictive relative if the immediate antecedent is not nominal?

(6) It's <u>in December</u> that she's coming

Another domain of variation concerns the discourse status of the sentence-final clause. The sentence-final clause in *it*-clefts is typically associated with expressing discourse-old information, shown in (7). In this example, we know from the prior text that a woman has been murdered by someone, so the open proposition *someone killed her* is given information. However, in other cases, such as (8), the underlined information is not given in the previous discourse and the proposition, that *someone*

once said 'laws are silent at times of war', does not even have to be known to the intended audience. Are these two functions related, or are they so divergent that we need to provide different sources for their origins?

(7) A: Is he the murderer?
B: No. It was the therapist that killed her

(8) (Start of lecture)
It was Cicero who once said, 'Laws are silent at times of war'

A further domain of variation concerns the relationship between specificational and predicational clefts. As I noted above, *it*-clefts have a specificational meaning. However, superficially similar proverbial sentences, such as (9), have a predicational meaning. For example, (9) is most closely paraphrased by the predicational copular sentence given in (10). In these examples, the postcopular element **describes** rather than **identifies** the referent. How does this structure relate to the specificational *it*-cleft, if at all? Can proverbial sentences, such as (9), really be called clefts? *It*-clefts also resemble extraposed sentences, such as (11). Do clefts share more than just an apparent likeness with this structure?

(9) It is a long road that has no turning
(10) A road that has no turning is a long one
(11) It is a miracle that he survived

This thesis provides answers to these questions by examining *it*-clefts within the framework of construction grammar (cf. Croft 2001; Fillmore, Kay and O'Connor 1988; Goldberg 1995, 2006; Kay and Fillmore 1999; Lakoff 1987). Construction grammar was developed with a view to providing full and explanatory accounts not only of broad generalizations but also of specialized linguistic patterns. In this model, constructions are not considered the epiphenomenal byproducts of a combination of componential meaning and highly general rules. Instead, aspects of form and meaning can be encoded

by the construction itself. Since much of the *it*-cleft's structure and use cannot be predicted from more general patterns of correspondence, it is well-suited to treatment within a constructional approach. I come back to this issue in §1.3.

In addition to providing a synchronic account, this thesis also examines *it*-clefts from a diachronic perspective. Relevant questions here include: What is the origin of the *it*-cleft? How did the different varieties of *it*-cleft emerge? Can historical evidence provide support for our synchronic analysis? And can an understanding of the history of English help to explain some of the highly specific properties of *it*-clefts? Most versions of construction grammar are usage-based theories and as such they are specifically designed to intersect with theories of acquisition, processing and language change (Goldberg 2006: 214). This study therefore provides a useful insight as to whether construction grammar and grammaticalization theory can be usefully integrated. I discuss this idea in more detail in §1.4.

This section has identified some of the phenomena that are the subject of this thesis and some of the questions that are addressed throughout. In the next section, §1.2, I provide an introductory background into the literature on cleft sentences. §1.3 asks why construction grammar is helpful in the treatment of *it*-clefts. Here, I present a brief outline of my own synchronic analysis and compare this to other constructional accounts proposed in the literature. In §1.4, I present some of the findings of previous historical studies of the *it*-cleft and my alternative diachronic construction grammar account is sketched in brief. §1.5 provides a note on the methodology employed in this study and an outline of the thesis is given in §1.6.

1.2 An overview of the literature on cleft sentences

As I noted in §1.1, authors tend to view *it*-clefts either from the perspective of their relationship to truth-conditionally equivalent subject-predicate sentences or from the perspective of their relationship to other specificational copular constructions. In this section, I outline these two approaches and the analyses that result from them. Although the individual proposals differ, these opposing viewpoints lead to two different kinds of analysis: those that treat the postcopular phrase as the preposed argument of the

proposition expressed in the sentence-final clause and those that consider the sentence-final clause to be associated in some way with the initial element *it*. The purpose of this section is not to provide an exhaustive and comprehensive review of the literature, but simply to highlight intellectual trends in the history of the analysis of the construction.

1.2.1 The expletive approach

For many authors, *it*-clefts are considered primarily as information structure variants of syntactically more basic sentences. From this perspective, *it*-clefts do not differ dramatically in their semantic content from canonical subject-predicate sentences, but are marked by the way that this informational content is presented (Ward, Birner and Huddleston 2002). Unlike their canonical counterparts, *it*-clefts have a fixed information structure; information that is to be foregrounded is placed in the postcopular focal position while the remaining semantic content is backgrounded into a sentence-final clause (see (12)).

(12) It was [[Howard]$_i$ [that ____$_i$ left]] [*it*-cleft]

(13) Howard left [canonical counterpart]

The analyses resulting from this approach assume that the focal element in *it*-clefts enters into a predication relationship with the information in the sentence-final clause; this accounts for their truth-conditional equivalence with simple subject-predicate sentences. From this, it follows that the initial element *it* and (in most accounts) the copular verb *be* are semantically empty, serving only to introduce, or foreground the postcopular element. As a result, in the cleft literature, these analyses are referred to cumulatively as the 'expletive' approach; common to all such accounts is the assumption that the initial pronoun *it* does not play an essential role in the interpretation of the sentence.

An early example of an expletive account is detailed by Jespersen (1937: 83-89). He suggests that *it*-clefts are syntactically identical to their noncopular counterparts except for the addition of a "lesser subject and verb" and a "connective word". So, for

instance, the elements *it*, *be* and *that* in (14) are semantically empty, with *Howard* and *left* entering into a predication relationship. This example is formalized, using Jespersen's notation, in (15).[2]

(14) [It was] Howard [that] left

(15) [sv] S [3^c] V

According to Jespersen, this analysis explains why *it*-clefts are used as a means of marking focus syntactically. He notes, "A cleaving of a sentence by means of *it is*...serves to single out one particular element of the sentence and very often, by directing attention to it and bringing it, as it were, into focus, to mark a contrast" (Jespersen: 1949: 147f).

A number of similar analyses were developed within the generative tradition of the 1980s. The details of these analyses differ. For example, Chomsky (1977) proposes that the postcopular element and the sentence-final clause represent a type of topicalization construction, while for Williams (1980) and Heggie (1988), these components are coindexed at the level of surface structure by a predication rule. Delahunty (1982, 1984) converts the sentence-final clause into a function (via lambda abstraction) which takes the postcopular element as its argument; after a reduction operation, the Logical Structure of cleft sentences is equivalent to that of their noncopular counterparts.

Although they use different mechanisms to accomplish it, these authors assume that *it*-clefts and truth-conditionally synonymous sentences must share a level of representation. Common to all of these accounts then is the treatment of the initial pronoun *it* as an expletive element, the analysis of the sentence-final clause as being in some way related to the postcopular element, and the understanding that the primary function of *it*-clefts is as a focusing device.

---

[2] In Jespersen's (1937: 86) analysis, the information in square brackets is not part of the sentence proper. Here, 3^c represents a "tertiary connective".

1.2.2 The extraposition approach

For others, *it*-clefts are considered foremost as specificational copular sentences. From this perspective, *it*-clefts are analysed in relation to corresponding pseudocleft sentences and sometimes to noncleft copular constructions with an identifying function. The term 'pseudocleft' is commonly used to encompass both *wh*-clefts, which are introduced by *wh*-words, and *th*-clefts, which are introduced by the definite article and one of a small number of semantically general head nouns such as *the one* or *the thing*.[3]

(16)	It's orange soda that I like best	[*it*-cleft]
(17)	What I like best is orange soda	[*wh*-cleft]
(18)	The thing that I like best is orange soda	[*th*-cleft]
(19)	My favourite drink is orange soda	[noncleft specificational]

Each of these examples has the function of identifying (or specifying) the postcopular focal element *orange soda*. Like the *it*-cleft in (16), the pseudoclefts also contain clausal elements. However, in (17) and (18) these clauses are in subject position. This suggests that the *it*-cleft is an extraposition construction: the sentence-final clause is not connected to the focal element; instead, it is related in some way to the initial *it*.

Again, an early example of such an approach is provided by Jespersen (1927). Prior to his 1937 account, outlined above, he proposed a "transposition analysis" of *it*-clefts. In the following passage, Jespersen suggests that *it*-clefts are paraphrased most closely by other specificational copular constructions. Here the sentence-final clause is analysed as a restrictive relative, modifying the constituent *it*. He notes, "...it is not really the antecedent (or what looks like the antecedent) that is restricted by a relative clause. When we say "it is the wife that decides" or "it was the Colonel I was looking for" what we mean is really "the wife is the deciding person" and "the Colonel was the man I was looking for": the relative clause thus might be said to belong rather to "it" than to the predicative following after "it is"" (Jespersen 1927: 88f).

[3] My use of the term '*th*-cleft' is from Collins (1991a, b). This term is used in a different sense in Ball (1977) and Hedberg (1990, 2000) to refer to cleft sentences introduced by demonstratives, such as *this* or *that*.

Other accounts that view *it*-clefts in relation to specificational copular sentences can be found in the transformational analyses of the 1970s, which derive *it*-clefts from pseudoclefts (see Akmajian 1970; Gundel 1977), or from the same source as pseudoclefts (see Wirth 1978). For these authors, the clausal element in subject position is extraposed, leading to the manifestation of *it* as either a placemarker or a pronominal copy. Again, the details of these analyses differ. For example, Akmajian and Wirth derive *it*-clefts via extraposition rules particular to cleft sentences, whereas Gundel suggests that this process is an instance of ordinary right-dislocation. For Gundel, the initial *it* is a pronominal copy of the right-dislocated clause, whereas for Akmajian, *it* is an expletive element.

Bolinger (1972) takes an approach that is more in line with Jespersen's (1927) original proposal, in which the relative clause restrictively modifies the constituent *it*. He suggests that analytic compound relatives, such as (20) "provide an ideal source" for *it*-clefts (Bolinger 1972: 110). Such sentences can undergo "inversion" (extraposition-from-NP), whereby the restrictive clause is extraposed but the nominal head remains *in situ*, as in (21). For *it*-clefts however, this "inversion" is obligatory rather than optional. In this analysis then, the constituent *it* is neither a placemarker, nor a pronominal copy, but a restrictively modified pronoun.

(20) That which he stole was money [analytic compound relative]
(21) That was money which (that) he stole ["inverted" compound relative]
(22) It was money which (that) he stole [*it*-cleft]

(examples from Bolinger 1972: 109)

Consequently, for these authors, there is little consistency as to the exact role of *it* or how the relationship between *it*-clefts and other specificational sentences works. Nevertheless, what these extraposition analyses share is a concern for recognizing *it*-clefts primarily as specificational copular sentences.

1.3 A constructional approach to clefts

So which of these approaches is the right one? From the perspective of construction grammar, the choice is straightforward. In this section, I explain why construction grammar is a useful framework for representing and accounting for the unusual properties of *it*-clefts and why an approach that examines clefts in relation to specificational sentences is more compatible with the principles of construction grammar. I go on to provide an overview of my own account of *it*-clefts before showing how this improves on the previous constructional analyses proposed in the literature.

Construction grammar is a non-derivational, monostratal model of language. In this model, linguistic patterns are understood as form-meaning pairs, much like individual lexical items. Since complex constructions are made up of smaller components, which are also form-meaning pairs, constructions can have compositional meaning. However, in construction grammar, these correspondences are internal to the construction and can therefore be construction-specific. In addition, the construction itself can encode meaning that cannot be attributed to its individual components. From this, it follows that construction grammar tolerates idiosyncratic linguistic patterns that cannot be predicted from highly general rules of the grammar. Within this framework, all aspects of form and meaning, including pragmatic and discourse-functional properties, make up a speaker's grammatical knowledge. Consequently, a constructional approach allows all of the *it*-cleft's properties, including idiosyncratic ones, to be represented in the grammar, facilitating a detailed and explanatory account.

The explanatory power of constructional accounts comes from the requirement that each construction must be motivated; that is, there must be some explanation provided for why this particular construction should exist in the language (Goldberg 2003: 120-1). For the most part, the motivation for a construction comes from within the grammar. In construction grammar, grammatical knowledge is represented as a network of constructions (form-meaning pairs). Specialized linguistic patterns inherit properties from more general patterns. The more properties a construction inherits, the more it can be said to be motivated by the language system. Constructions that are related to one another are shown to inherit properties from the same general pattern, forming a

"family" of constructions. A more detailed introduction to construction grammar is provided in chapter 2.

The organization of grammatical knowledge in construction grammar predicts that analyses of *it*-clefts based upon their relationship to noncopular subject-predicate sentences will have less explanatory adequacy than an approach that views *it*-clefts in relation to other specificational copular sentences. Goldberg (1995: 108) says that inheritance links are only posited between constructions that are formally related in terms of surface structure: "The intuition is that the existence of a given form with a particular meaning in no way motivates the existence of a different form with a closely related meaning". Consequently, although noncopular sentences can often be used to paraphrase *it*-clefts, this truth-conditional synonymy is not expressed in the grammatical system. As predicted, accounts that view *it*-clefts in relation to structurally less complex sentences leave a number of questions unresolved: Why should focus be marked using this particular structure? Why do *it*-clefts have so many semantically empty elements? Where do the existential and exhaustiveness presuppositions come from?

In this thesis, I view the *it*-cleft foremost as a member of the family of specificational copular constructions. *It*-clefts, *wh*-clefts, *th*-clefts and noncleft copular sentences all inherit properties from a more general, schematic, specificational copular construction. But what is a specificational copular construction? And where does specificational meaning come from? Unfortunately, the answers to these questions are not obvious and a number of different analyses have been proposed in the literature. In order to understand the larger schema, or category, of copular constructions, I provide my own analysis. I argue that in specificational copular constructions, specificational meaning results from a special kind of nominal predication involving definite noun phrases. I treat nominal predication set theoretically, as a semantic relation between members and sets. I conclude that this analysis has a number of advantages over alternative accounts of specificational sentences and is more successful at explaining the data.

My analysis shows how a number of the *it*-cleft's properties are motivated. Like the extraposition accounts of Jespersen (1927) and Bolinger (1972), I analyse the

sentence final clause as a restrictive relative, modifying the initial *it*. In line with the proposals of Hedberg (1990, 2000), Percus (1997) and Han and Hedberg (2008), I argue that *it* and the relative clause together operate like a discontinuous definite description. Since definite descriptions contain exhaustiveness and existential presuppositions, this analysis explains why these properties are found in *it*-clefts. It also reduces the number of semantically "dummy" elements. For example, initial *it* is shown to perform an important quantifying function. However, where my own analysis represents an improvement on the literature is in providing an explanation as to why definite descriptions are crucial for creating the specificational meaning of *it*-clefts as well as other copular constructions. As I explain in chapter 5, the analysis of specificational sentences that I advance allows for a straightforward account of the relationship between specificational and predicational/proverbial *it*-clefts.

Alternative constructional accounts of *it*-clefts have been proposed by Lambrecht (2001) and Davidse (2000). However, neither of these authors makes use of a system of inheritance. Consequently, both present highly idiosyncratic analyses. For example, Lambrecht (2001) examines *it*-clefts in relation to noncopular subject-predicate sentences. As a result, his analysis suffers from the same lack of explanatory adequacy as other expletive accounts. Davidse (2000), on the other hand, analyses the *it*-cleft as a highly complex structure involving two clauses (one of which is unique to cleft constructions) which enter into different semantic relationships with the postcopular element. Again, this structure is not shown to be motivated by the language system.

The constructional approach provided in this thesis is therefore superior to the accounts of Lambrecht (2001) and Davidse (2000) since it makes full use of the tools exploited in construction grammar for making **generalizations**. By examining *it*-clefts in relation to the taxonomy of specificational sentences and exploiting an appropriate inheritance hierarchy, I maximize the motivation for the *it*-cleft construction. Only after examining *it*-clefts in relation to the rest of the grammar are the exceptional or truly construction-specific characteristics isolated. As Goldberg (2003: 118) observes, "a given construction often shares a great deal with other constructions that exist in a language; only certain aspects of its form and function are unaccounted for by other

constructions". As a result, the constructional approach advocated in this thesis has much greater explanatory power.

1.4 The grammaticalization of the cleft construction

We have seen then that construction grammar tolerates (but nevertheless seeks to limit) idiosyncrasies in the language system. However, ideally, even exceptional properties should be provided with an explanation of some sort. According to Goldberg (2003: 121), in such cases, motivation can be provided by factors external to the grammar. In this section, I ask whether historical evidence can provide motivation for, and therefore explain, the construction-specific properties of the *it*-cleft.

From my synchronic analysis of the *it*-cleft as a type of specificational copular construction, certain structural aspects remain a puzzle, such as the modification of *it* by a restrictive relative clause and the extraposition of the relative clause. By exploring the history of the English language, I show that although these properties are no longer motivated by the language system, they are likely to have been inherited from formally related constructions existing at earlier periods of the language. In this way, the *it*-cleft shows how the retention or 'entrenchment' of once-motivated form-function pairings can lead to construction-specific properties which are no longer productive in other areas of the grammar. For most types of *it*-cleft then, the seemingly idiosyncratic structure can be shown to be motivated by the language system at least at the point of origin. However, there are subtypes of *it*-cleft which exhibit properties that cannot be attributed to inheritance at any period of the language.

In §1.1, I introduced two types of *it*-cleft which are sometimes treated as separate constructions from the *it*-cleft prototype: those with non-nominal foci and those with new information in the sentence-final clause. The particular range of *it*-cleft foci is not shared by other specificational copular constructions. Likewise, other kinds of specificational sentence cannot express totally new, but factually presupposed information. For example, the *th*-cleft in (24) seems strange when occurring in discourse-initial position and suggests that we should already know that *someone once said 'laws are silent at times of war'*.

(23) (Start of lecture)

It was Cicero who once said, 'Laws are silent at times of war'.

(24) (Start of lecture)

#The one who once said 'Laws are silent at times of war' was Cicero.

As a result, the range of non-nominal foci and the ability to express hearer-new information are properties which are not inherited from the wider specificational construction. This begs the question, where did these more idiosyncratic types of *it*-cleft come from? Do they have a separate source from *it*-clefts with nominal foci and given information in the relative clause? Are they in fact separate constructions from the *it*-cleft proper?

In a previous historical study, Ball (1991, 1994a) found that *it*-clefts with non-nominal foci and examples with new information in the relative clause originated at a later date than NP-focus *it*-clefts expressing discourse-old information. Although Ball's study is exclusively diachronic, she emphasizes present-day structural and functional differences between *it*-clefts with NP and non-NP foci and between *it*-clefts with given and new information in the relative clause. As a result, Ball argues that these different types of *it*-cleft derive from separate sources. She suggests that over time the existing *it*-cleft tokens merged with an increasing number of different configurations to form a structurally distinct 'non-NP focus *it*-cleft' and the 'informative-presupposition (IP) *it*-cleft'.

In contrast, my own synchronic analysis of *it*-clefts can incorporate instances with different focal categories and with either given or new information in the relative clause; these do not require a distinct structure and their functions are shown to be related. Using data from the Penn Parsed Corpora of Historical English, I show that the less-prototypical *it*-cleft subtypes have emerged gradually via extension from the existing *it*-cleft prototype. These emerging subtypes involve mismatch phenomena: that is, they contain construction-specific mappings of form and meaning. The development of the *it*-cleft therefore involves a change from a configuration which is fully motivated by inheritance from more general patterns of correspondence, to one that encodes

meaning which is no longer necessarily predictable from the meanings associated with its component parts. As Langacker (1991: 295) comments, such "extensions from the prototype" occur because of the "pressure of adapting a limited inventory of conventional units [or constructions] to the unending, ever-varying parade of situations requiring linguistic expression".

In addition to showing how general principles of language change can provide motivation for the more idiosyncratic properties of the *it*-cleft construction, I also discuss whether the construction grammar framework is useful for theories of language change. I examine the development of the *it*-cleft in relation to grammaticalization theory. Although grammaticalization is usually said to apply to individual lexical items which become more grammatical over time, I investigate whether changes to more complex constructions, such as the *it*-cleft, exhibit properties consistent with grammaticalization, such as gradualness and unidirectionality. I ask whether grammaticalization theory should be extended to larger constructions and whether the construction grammar model of grammatical knowledge could change grammaticalization theory for the better.

1.5 Methodology

This thesis examines the English *it*-cleft from both a synchronic and a diachronic perspective. My analysis begins by integrating *it*-clefts into an original account of specificational copular constructions. In these chapters, I rely largely on examples that are either invented or taken from the literature. I have chosen to exemplify my discussion in this way for several reasons. First, I am engaging with a literature where the use of invented examples is common practice. Also, since the issues surrounding the data are often complex, I have made an effort to keep examples brief and to choose examples that highlight the relevant features without requiring unnecessary explication. In this part of the thesis, the focus of my discussion is on the prototypical *it*-cleft subtype. As a result, I am not interested here in detailing qualitative differences between individual instances. Instead, the data are intended purely to exemplify the discussion. My choices in response to these issues mean that although this thesis adopts a usage-

based construction grammar framework, there is perhaps less actually occurring data than in similar studies of this type.

My analysis concludes with an examination of the historical development of the *it*-cleft construction, in which I make use of data from the Penn Parsed Corpora of Historical English. My investigation focuses on the Late Middle and Early Modern periods of English. The Late Middle English data is taken from the Penn-Helsinki Parsed Corpus of Middle English second edition (PPCME2). From this corpus, I extracted a subcorpus of 894341 words containing texts composed between 1300 and 1500 with which to conduct my search. The Early Modern data comprises all 1794010 words of the Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME) which contains texts written between 1500 and 1710. Both of these corpora were searched using the specially designed CorpusSearch2 program. Since these corpora are from the same series, they are comparable in terms of syntactic annotation, size of the corpus, size of text samples and range of genres. In addition to my historical data, I also discuss some of the more interesting present-day English examples found in the British component of the International Corpus of English (ICE-GB).

I use these data both to elucidate my discussion and to provide qualitative and quantitative evidence of diachronic change. However, since *it*-clefts are not a high frequency construction and the historical corpus data is limited, it is to be understood that quantitative data is not always reliable. As a result, I ensure that my analyses are supported by qualitative data and by well-attested theories of language change.

I undertake a corpus-based rather than a corpus-driven study; my synchronic analysis informs my diachronic investigation and, in turn, the diachronic evidence is used to support my synchronic account of *it*-clefts.[4] One of my reasons for approaching corpus study in this way is that *it*-clefts are notoriously difficult to identify and separate from superficially similar but functionally distinct examples, such as extraposed sentences (see Haugland 1993). As a result, it is important to understand these often subtle differences before extracting the relevant data. Furthermore, since there are so

[4] See Tognini-Bonelli (2001) for an outline of the differences between these two approaches to corpus study. While I advocate a corpus-based approach in this thesis, Tognini-Bonelli discusses the merits of a corpus-driven approach.

many different analyses of *it*-clefts, linguists may differ over which examples count as clefts. Regardless of whether a linguist has fully developed, or settled on, an analysis, they inevitably bring such assumptions with them when they come to sorting through the data. I believe that it is impossible to be completely theory neutral and as a result I am of the opinion that it is better to know what your criteria are outright and from the very beginning, so that at least the approach is consistent. This also means that I can address the issue of how well my theory accounts for the data.

1.6 An outline of the structure of the thesis

In the present chapter, I have sketched a brief overview of the issues that this thesis addresses, provided some introductory background material and summarized the main arguments that I propose. In the following chapter, I lay out the theoretical model of construction grammar, the mechanisms which I make use of in my account and the theoretical assumptions on which the rest of the thesis is based.

My analysis begins in chapter 3 with a justification of the position that *it*-clefts should be examined in relation to other specificational constructions. After finding various faults with the existing accounts of specificational copular sentences, I provide an original, constructional analysis. My argument is situated as a response to the 'inverse' accounts formulated within the minimalist tradition. While I agree that specificational sentences involve mismatch, whereby the syntactic subject functions as the semantic predicate, I do not assume that specificational meaning is the product of syntactic movement. Instead, taking a set-theoretic perspective, I claim that specificational meaning results from a special type of nominal predication relationship involving definite noun phrases.

In chapter 4, I integrate *it*-clefts into my analysis of specificational copular constructions. I explain that motivation for the *it*-cleft is maximized if we adopt a (non-derivational) extraposition-from-NP analysis of *it*-clefts in which the initial *it* and the sentence-final clause operate together like a discontinuous definite description. From this perspective, many of the seemingly idiosyncratic properties of the *it*-cleft are shown to be inherited from the more schematic and general specificational copular

construction. Together, chapters 3 and 4 represent the most fundamental components of this thesis, since many of the arguments presented in the remaining chapters arise out of the analyses of specificational copular sentences and *it*-clefts proposed here.

In chapter 5, I examine whether my *it*-cleft analysis can be extended to accommodate all of the different types of *it*-cleft outlined in §1.1, including predicational and proverbial clefts, *it*-clefts with non-nominal foci and informative-presupposition (IP) *it*-clefts. I show that from my analysis of specificational sentences as involving a special type of nominal predication, the relationship between specificational and predicate nominal sentences, and consequently the relationship between specificational and predicational (including proverbial) *it*-clefts, becomes straightforward. This represents an important advantage to my analysis. For non-NP *it*-clefts and IP *it*-clefts, on the other hand, I show that while the analysis of *it*-clefts outlined in chapter 4 can accommodate these subtypes into a unified analysis, they contain construction-specific properties which are not inherited from more basic constructions. I suggest that a historical investigation into the *it*-cleft construction may help us to understand where these idiosyncratic properties came from. I come back to this issue in chapter 7.

My synchronic analysis ends with a discussion about the constructional framework I adopt. In chapter 6, I outline what is constructional about my account of *it*-clefts and show how it compares favourably to the other so-called constructional analyses proposed in the literature. Finally, after explaining what the advantages are to examining the *it*-cleft within the framework of construction grammar, I provide a constructional taxonomy of cleft and specificational sentences based upon the inheritance relations posited so far.

Chapters 7 and 8 comprise a historical investigation into the *it*-cleft construction. In chapter 7, I make use of historical evidence to provide an explanation for the *it*-cleft's idiosyncratic characteristics. I show that many of the *it*-cleft's construction-specific structural properties are remnants of once regular patterns in the history of English. I conclude that these properties were originally inherited from constructions which no longer exist in the present-day language system. Using corpus data, I go on to show how

it-clefts with non-nominal foci and IP *it*-clefts have emerged via extension from the existing *it*-cleft prototype, according to general principles of constructional change. My analysis of the origin and subsequent development of the English *it*-cleft contrasts with Ball's (1991, 1994a) account. As I explain, there are a number of problems with Ball's argumentation.

In chapter 8, I show how the diachronic data provides a useful insight into the nature of the present-day *it*-cleft construction and explains the subtle ways in which it differs from other types of specificational copular sentence. As part of this discussion, I ask what the historical development of the *it*-cleft tells us about the nature of constructional change and whether grammaticalization theory is applicable to changes involving complex constructions rather than atomic lexical items. I conclude this chapter by showing how the historical evidence can be used to both inform and support my synchronic analysis of the present-day *it*-cleft, resulting in a maximally explanatory account. A summary of the thesis is given in chapter 9 along with my final conclusions.

2. CONSTRUCTION GRAMMAR AS A MODEL OF LANGUAGE STRUCTURE AND LANGUAGE CHANGE

2.1 Introduction

In this chapter, I introduce the principles and concepts that are specific to a usage-based, constructional theory of grammar. The purpose of this chapter is to outline some of the basic claims that underlie this thesis and to explain the machinery that I make use of in my analysis of cleft sentences. In this section, I provide some historical background into why the construction grammar framework was developed and what it was hoped to achieve. I also explain the principles behind the usage-based model of change that I adopt. In addition, I distinguish between the different variants of construction grammar and explain which set of theoretical assumptions I adopt in this thesis and why.

Construction grammar was developed as an alternative to the componential model of grammatical knowledge proposed by theories of generative grammar (see Croft and Cruse 2004: 225). In a componential model, each type of linguistic knowledge (syntax, semantics and so on) makes up a separate component. Aspects of meaning and form are mapped on to one another by general linking rules. The only idiosyncratic and item-specific mappings between these components are found in lexical items. Croft and Cruse (2004: 227) illustrate this model with the diagram given here as Figure 2.1. Note that the lexicon is represented as the only vertical component, combining information from the other horizontal components. Constructions do not have theoretical status in this model; they are purely epiphenomenal products of componential meaning and general rules of the grammar (see Chomsky 1995: 170).

The way that the grammar is organized in the componential model suits the generative theory that syntax can be studied independently from semantics and other aspects of meaning, such as pragmatics and discourse function. As a consequence of this assumption, "functional differences between formal patterns [are] largely ignored" (Goldberg 2006: 4). Within generative grammar, constructions are simply syntactic configurations; the fact that constructions may encode noncomponential meaning is either not recognized or is deemed to be outside of the scope of the "core" phenomena

that are regarded as the focus of linguistic study. This notion of "core" language arises from the generative assumption that such phenomena result from innately programmed linguistic knowledge, forming part of Universal Grammar. Other linguistic phenomena, such as discourse-functional properties and (oftentimes) pragmatic properties, are learned inductively from experience. As such they are relegated to the "periphery" of language and are not acknowledged as the focus of linguistic study.

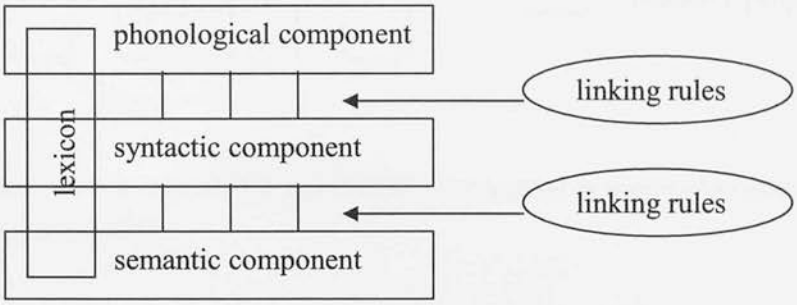


Figure 2.1 *The organization of grammatical knowledge in a componential model (Croft and Cruse 2004: 227)*

For construction grammarians, on the other hand, all aspects of language are the proper objects of linguistic study. In this framework, constructions are given a theoretical status as symbolic form-meaning pairings, much like lexical items. Unlike in componential models, which map separate components of linguistic knowledge onto one another by general linking rules, in construction grammar, the form-function mapping is represented as internal to the construction. From this, it follows that that syntax cannot be studied in isolation and all aspects of meaning, including semantics, pragmatics and discourse-function are required to give a full account of grammatical knowledge. This model of language structure therefore anticipates, rather than ignores, construction-specific information, which is not predictable from more general patterns of correspondence. For construction grammarians then, the aim is to provide full and explanatory accounts of these specialized linguistic patterns. Croft and Cruse (2004:

256) represent the model of grammatical knowledge in construction grammar with the diagram given here as Figure 2.2.

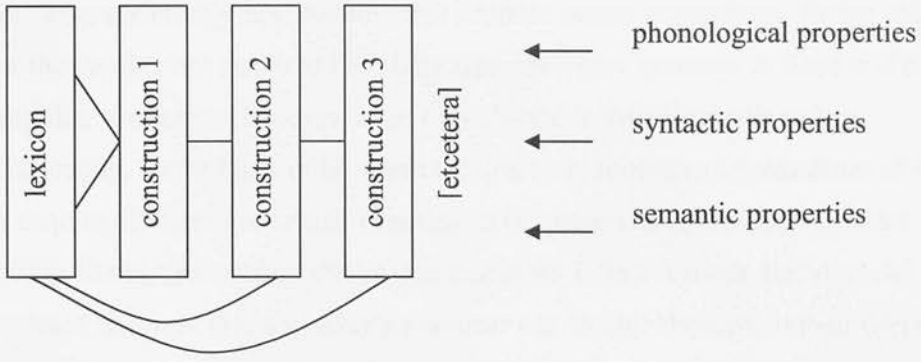


Figure 2.2 *The organization of grammatical knowledge in construction grammar (Croft and Cruse 2004: 256)*

Most variants of construction grammar are usage-based frameworks. The usage-based model claims that all grammatical knowledge, not just the periphery of language, is learned inductively from the input. In this model, it is not assumed that language users are innately programmed with linguistic knowledge; instead the assumption is that language learning involves general cognitive processes. From this it follows that all of the grammar, including both specialized linguistic patterns and broad generalizations, should be given a uniform representation. As a result, in usage-based constructional theories, all of the grammar is represented as constructions (form-meaning pairs). Linguistic patterns are represented as constructions in the language system if either some aspect of their structure or use is unpredictable or their instances are sufficiently frequent that the speaker is likely to induce an abstract mental schema. Goldberg (2006: 18) says that "the network of constructions captures our grammatical knowledge of language *in toto*, i.e. **it's constructions all the way down**" (italics and emphasis original).

An important advantage to a usage-based theory of construction grammar is that it is able to intersect with and work alongside theories of language change. In this thesis, I examine the extent to which a construction grammar model of grammatical knowledge

can be usefully integrated into grammaticalization theory. Here, I briefly outline some of the principles behind a usage-based model of change.

In theories of grammar which assume that core language is hardwired in the brain, language change is something that happens across generations, during acquisition. After the speaker has acquired their language, their core grammar is fixed and is not susceptible to change. However, what may change is their linguistic output. Consequently, on the basis of the input they receive, subsequent generations of speakers may acquire different grammars, resulting in language change. In such a model, language change is therefore abrupt (see Lightfoot 1979). A usage-based model, on the other hand, assumes that a speaker's grammar can change throughout their lifetime. In this model, language learning and language change involves inductively generalizing over instances to form schemas which are represented in the language system. Since this involves general cognitive processes, such as categorization, there is no need to assume that change only occurs in the stages of acquisition. In such a model, changes to the grammar are not catastrophic, but occur in incremental stages, giving the impression of gradualness. In this thesis, I assume a usage-based model of language change.

However, not all versions of construction grammar are usage-based. Goldberg (2006) identifies four main theories of construction grammar: Unification Construction Grammar (Fillmore, Kay and O'Connor 1988; Fillmore *et al.* forthcoming, Kay and Fillmore 1999), Radical Construction Grammar (Croft 2001), Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995, 2006) and Cognitive Grammar (Langacker 1987, 1991). Of these theories, Unification Construction Grammar has developed somewhat separately; unlike the others, it is not a usage-based theory. Throughout this chapter, I indicate where Unification Construction Grammar differs from the other constructional theories. I do not follow the theoretical assumptions of Unification Construction Grammar in this thesis. Instead, my approach is most consistent with the theories presented by Goldberg (2006) and Croft (2001). Although Croft's focus is on creating a theory suitable for cross-linguistic typological study, both authors agree on the same set of fundamental principles (Goldberg 2006: 220-226).

In the following sections, I introduce the relevant concepts and machinery of construction grammar that I make reference to throughout the thesis. These include: the inheritance hierarchy as the organization of grammar, the construction as a symbolic form-meaning pair, mismatch and coercion, noncompositionality and construction-specificity. For each concept, I explain how they are relevant to a usage-based theory of language change and show how they interface with the basic claims of grammaticalization theory.

2.2 Classification and inheritance as a formal system and as a model of change

In construction grammar, it is assumed that constructions form a structured inventory which makes up the speaker's knowledge of the language. This inventory is represented as a taxonomic network of constructions with each construction constituting a separate node (Croft and Cruise 2004: 262). The network is hierarchical, showing that some constructions are more basic or general than others. Lower-level constructions inherit attributes from higher-level constructions.

Individual construction grammars differ as to which model of inheritance they assume. The complete mode of inheritance is typically adopted in unification-based grammars, such as Kay and Fillmore (1999). In this model, a construction inherits all of the information that is specific to a dominating construction. This relation allows information to be represented only once in an inheritance hierarchy. Inherited information is stored only in the dominating construction, at the highest level possible. Complete inheritance therefore licenses a non-redundant system of linguistic knowledge. This means that for a speaker to classify an utterance as an instance of a lower-level construction, they must search all the way up the inheritance tree in order to determine what the specifications of this construction are and whether the utterance matches them.

From the complete inheritance model, it follows that although lower-level constructions may contain information that is not present in the more general, dominant construction, they may not contain information that conflicts with information presented at a higher level. To do so would result in ill-formedness. As a result, categories formed

on the basis of these inheritance hierarchies are classical. That is, a construction is defined as a member of a larger constructional category if and only if it inherits all of the grammatical structure of the superordinate construction.

Other constructional theories adopt default inheritance; these include Cognitive Construction Grammar (Goldberg 1995, 2006; Lakoff 1987), Radical Construction Grammar (Croft 2001), Cognitive Grammar (Langacker 1987, 1991) and also Word Grammar (Hudson 1990, 2007).[1] A crucial property of the default inheritance model is that conflict between the information specified in inheriting constructions and information specified in dominating constructions is permitted. This means that all of the attributes of a dominating higher-level construction will be inherited by the lower-level construction **unless** there is conflict. In this case, the more specific construction "wins out" and inheritance is limited to only non-conflicting information.

In default inheritance models, information is often stored redundantly. A redundant system represents information not only on the highest possible node, but at all levels in the hierarchy. Goldberg (1995: 73-74) argues in favour of a redundant, or "full entry", model since without it there is no way of resolving conflict that may arise in cases of multiple inheritance. For example, if a construction inherits information from two different dominating constructions, but the information specified for these constructions conflicts, how do we know which of these parent constructions "wins"? In a case such as this, we can only know which information is inherited from which parent construction if this is specified redundantly in the daughter construction. A redundant system can also be argued for on psychological grounds; although a non-redundant system is more economical, it requires maximum on-line processing. Goldberg (1995: 74) says that "the inheritance mechanism of our system is not an on-line process, but rather a static relation defined by shared information…"

---

[1] Word Grammar shares a great deal with constructional theories of grammar. For instance, in this theory, grammatical knowledge is represented as a network, with inheritance relations connecting the individual nodes. Hudson (2007: 153) also recognizes that the syntax of language is comprised "of a very large number of constructions, each with its own peculiar interactions with other constructions and with lexical items". The main difference between Word Grammar and other constructional theories is that the former focuses on the (more or less specific types of) dependency relations between the words which make up these constructions (Hudson 2007: 156).

This thesis makes use of default inheritance and assumes that information is stored redundantly. Unlike alternative constructional accounts of cleft sentences, I place *it*-clefts in a constructional hierarchy of copular sentences. I argue that my analysis of *it*-clefts is superior, since a large number of the *it*-cleft's properties can be attributed to inheritance from the more basic specificational copular construction. As Goldberg (2003: 120) comments, "What imbues a constructional approach with explanatory adequacy is a further desideratum that each construction must be *motivated*" (italics original). The more properties that are inherited from other constructions in the language, the more we can say that the construction is motivated. The explanatory power of motivation is in some way reliant on a redundant system. As Lakoff (1987) notes, the more redundancy a construction exhibits, the better it fits into the linguistic system. An analysis that maximizes motivation can therefore explain why this construction with these particular form-meaning correspondences should be likely to occur in the language.

Default inheritance is designed to allow for partial generalizations to be recognized. Some constructions or instances of constructions are "better" (or more motivated) members of the constructional category than others. As a consequence of this, categories defined by the inheritance hierarchies of this model are non-classical, with each category containing a prototypical member (or members) and non-prototypical members.

In a usage-based model, the storage and organization of grammatical knowledge is dependent upon, and can change according to, patterns of activation. Together, non-classical categories and usage-based assumptions make a number of testable hypotheses regarding the diachronic development of constructions. Non-prototypical instances of a constructional category are formed by extension from the prototype, overriding inheritance from the overarching construction. As prototypical and non-prototypical instances coexist, the speaker forms an inductive generalization (or abstraction) which stipulates only those characteristics which are shared by all of its members. This is consistent with the usage-based generalization that the entrenchment of a more abstract (or more general) schema in the speaker's inventory is a function of type frequency; that

is the frequency of different "types" of instance. Changes to the conceptualization of the overarching category may, in turn, have consequences for yet higher-order constructions (or categories) in the taxonomy. This type of constructional change therefore proceeds upwards throughout the hierarchy, leading to the creation of new constructions and the reconfiguration of existing ones. As Goldberg (2006: 62) comments, "...we constantly parcel out meaning, form abstractions, and generalize over the instances we hear".

However, the usage-based model predicts that not all constructional change will follow this same pattern. While type frequency results in the entrenchment of a more abstract schema, token frequency (that is, the number of times a given instance is activated) results in the entrenchment of the instance. Bybee (1985: 132-134) argues that the repeated use of an instance which is stored as a conventional unit in the speaker's grammar will only activate (and strengthen) the entrenched instance; it will not serve to reinforce the superordinate schema. Consequently, the entrenchment (or fossilization) of individual instances is often found occurring alongside a concomitant loss (or weakening) of the overarching schema, which is no longer type productive.

A usage-based, constructional approach therefore predicts that there are two different types of constructional change: one which is brought about by type frequency and one which is dependent upon token frequency. Although these changes have consequences for constructions at different levels in the hierarchical network, they nevertheless involve the same process of conventionalization (or the entrenchment of schemas). For historical linguists who accept a constructional model of language structure, these two different types of constructional change constitute a reimagining of the well-attested item-based changes of grammaticalization and lexicalization, respectively (see especially Himmelmann 2004 and Trousdale 2008a). As I explain in chapters 7 and 8 of this thesis, the diachronic development of the *it*-cleft construction (including changes to its dominating constructions) involves both 'schematization' and 'fossilization'; that is, both grammatical and lexical 'constructionalization'.

2.3 Signs as the basis of grammar

In componential syntactic theories, in which separate components of the grammar interrelate via linking rules, constructions are epiphenomenal. However, in construction grammar, the syntactic elements and semantic components that are particular to each construction are related by symbolic links that are internal to the construction. In this way, each complex construction contains units of form-meaning pairs. However, as well as linking individual elements to their conventional meaning, there is an additional symbolic link that relates the entirety of the construction's form to the construction's conventional meaning. As a result, this model can represent aspects of conventional meaning that are associated with the construction, but cannot be attributed to any of its individual formal elements. Consequently, even a complex construction is itself a symbolic whole, or to use Saussure's term, a linguistic sign. Croft and Cruse (2004: 258) represent the symbolic structure of a construction using the illustration given here as Figure 2.3. As shown in this diagram, the construction's formal characteristics are made up of syntactic, morphological and phonological properties, while its conventional meaning comprises semantic, pragmatic and discourse-functional properties.

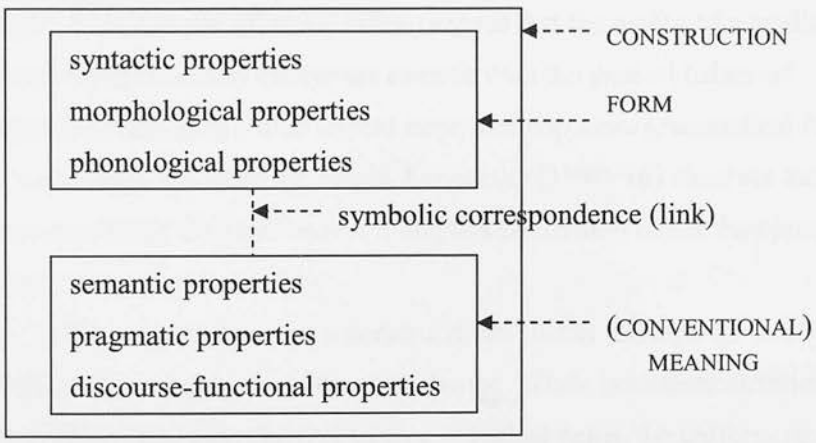


Figure 2.3 *The symbolic structure of a construction (Croft and Cruse 2004: 258)*

In unification-based construction grammars, a linguistic pattern is only considered a construction if some aspect of its form or function is not predictable either

from its component parts or from other constructions that exist in the language. Such patterns must be listed as constructions in order to provide a comprehensive account of their properties. However, in usage-based construction grammars, predictable or regular patterns of the language can also be recognized as constructions; as long as they occur with enough frequency, they may be stored as symbolic units (Goldberg 2006: 05). In a usage-based model then, constructions are simply conventionalized chunks of linguistic knowledge.

As symbolic units, constructions are very much like lexical items. The only difference is that while constructions are complex, words are simple and may even be atomic (that is, morphologically simple). In construction grammar then, the lexicon and the inventory of grammatical constructions are not separate components and instead exist along a 'syntax-lexicon' continuum ranging from the most schematic (or general) and complex constructions to fully substantive and simple items. Construction grammar therefore allows a uniform representation of all grammatical knowledge, including words and morphemes, as comprising an inventory of signs.

According to some authors, the central role of the sign in construction grammar makes it ideally suited for providing a theory from which to study grammaticalization changes. For example, Diewald (2006) argues that the notion of a gradient continuum between the lexicon and the syntax accords with the gradual nature of grammaticalization, in which lexical items develop more grammatical functions through a series of small intermediate stages. Langacker (1990: 16) observes that a grammaticalizing element "moves along this continuum rather than jumping from one discrete component to another".

However, if we accept a constructional model of language structure, then our definition of grammaticalization must change. While grammaticalization theory traditionally focuses on changes to atomic lexical items, the uniform representation of both lexical elements and complex constructions as signs suggests that changes which apply to substantive, simple constructions should also affect more schematic and complex constructions. In other words, if lexical items can grammaticalize, then larger,

less substantive constructions should also be subject to grammaticalization (Trousdale 2008b: 33-34).

2.4 Mismatch and coercion

The construction grammar model of grammatical knowledge as comprising a structured inventory of symbolic units encourages the linguist to think about form and meaning at the same time. This, in turn, allows the construction grammarian to identify and examine mismatch phenomena. Francis and Michaelis (2003: 2) use the term 'mismatch' to describe mappings between form and meaning which do not conform "to more general patterns of correspondence in the language". They identify two different kinds of mismatch phenomena. 'Complexity mismatch' occurs when there is not a one-to-one relationship between formal elements and semantic components. For example, extraposition constructions, such as (1) below, involve an expletive *it* which is "present in syntax but semantically unspecified" (Francis and Michaelis 2003: 4).

(1) **It** is a miracle that he survived

'Content mismatch', on the other hand, involves an incongruous mapping between form and function. For example, Francis and Michaelis provide the example of predicate nominals. Although noun phrases prototypically function as referential argument expressions, in sentences like (2) they function as predicates. This construction therefore involves category mismatch, whereby "the typical formal properties of one lexical category…are associated with the typical semantic properties of another category" (Francis and Michaelis 2003: 5).

(2) John is **a doctor**

In derivational theories, mismatch is accounted for by assuming that there is an underlying level in which the structure behaves in the expected ways. The mismatch configuration is derived from default patterns of correspondence using transformations

or movement operations. Such an approach can account for the complexity mismatch in extraposed sentences, for example. On this analysis, the sentence in (1) derives from the underlying non-extraposed sentence in (3).

(3) That he survived is a miracle

However, cases of category mismatch (such as predicate nominals) are problematic for these syntax-based theories, which assume that semantic distinctions are represented by syntactic categories and constituent structure. As a result, some authors, including Radford (1997), have suggested that predicate nominals represent a separate phrase structure category from noun phrases serving as arguments, making use of the terminological distinction between noun phrase (NP) and determiner phrase (DP).

In contrast, for construction grammars which allow default inheritance, mismatch can be accommodated neatly, without needing to invoke any additional mechanisms. Since mapping between form and meaning is internal to the construction in this model, mismatch constructions are simply presented as containing information that overrides inheritance from more general patterns. As Francis and Michaelis (2003: 24) comment, mismatch effects therefore "provide evidence for symbolic constructions and for inheritance hierarchies". While derivational theories have limits on length of movement and possible landing sites, construction grammar constrains degrees of mismatch with the condition that every construction must be motivated, inheriting a maximum number of properties from the language system (see §2.2).

Mismatch often occurs as the result of language change, particularly grammaticalization. For example, Traugott (2007) discusses a range of examples, such as *a lot of* and *a bit of*, which underwent a development from partitive to degree modifier. As partitives, these constructions had meanings similar to *a part of* or *a share of*. Since a part of something suggests a quantity, these partitives were associated with quantifiers via pragmatic implicature. For example, *a bit of* derives from *a bite out of* and consequently implies a small piece or quantity. This enabled a semantic reanalysis from partitive to degree modifier in which the head develops a scalar meaning, resulting

in constructional polysemy and extension to nominal contexts in which a partitive meaning is unavailable, such as *a lot of folks*.

Semantic reanalysis created a partial mismatch between form and meaning; at this point, these constructions have the syntactic attributes of partitives, but the semantics of quantifiers. Over time, these constructions co-occurred with fewer types of determiner (**the lot of*) and acquired a less compositional structure, as shown by the loss of integrity in *sorta* and *alotta*. As a result, these binominal strings were realigned with the degree-modifier category, resulting in syntactic reanalysis (or 'head shift') from [NP1 [of NP2]] to [[NP1 of] NP2]. The constructions then developed adverbial functions and extended distributions, modifying adjectives as in *a lot wiser* and *sort of cold*. Such developments brought about subsequent changes for the overarching degree modifier construction, which became a more schematic and productive category.

In this example of grammaticalization (from partitives to degree modifiers), mismatch is the byproduct of a development in which certain constructions become integrated into a different category or family of constructions. Only once they have become full members of this category, is mismatch, to a large degree, resolved (see Francis and Yuasa 2008). However, other times, mismatch is not associated with wholesale category change; in such cases, mismatch may become a conventionalized feature of the language system.

For example, complex constructions may be extended to accommodate items with which they are not prototypically associated. Michaelis (2003: 263) uses the terms 'coercion' and 'coercion effect' "for the enriched interpretations which result from this procedure". In construction grammar, coercion is accounted for by a combination of default inheritance and the symbolic nature of the construction. If the construction's conventional meaning conflicts with the meaning typically associated with a superimposed lexical item, then the constructional requirements "win out" and the lexical item conforms to them. Michaelis (2003: 268) refers to this as the 'Override Principle'.

Alternatively, in modular theories of grammar, which separate syntactic and semantic levels of representation, coercion effects are brought about by placing specific

coercion operators in semantic structure (see Jackendoff 1997 and De Swart 1998). These operators are triggered by the need for functors to receive suitable arguments. However, as Michaelis (2003: 263-264) comments, while this method can account for mismatches such as *a beer*, in which the indefinite article requires an operator to derive a count noun from a mass noun, it cannot account for template-based coercion. In these cases, the syntactic head cannot be interpreted as triggering coercion. Instead, the construction as a whole alters what the word designates.

An example of template-based coercion is provided by Goldberg (1995: 158), who identifies cases of mismatch in examples of the English caused-motion construction between the semantics of the verb and the semantics designated by the construction. She notes that in examples like *Joe kicked the dog into the bathroom*, motion is coded by the verb and the preposition *into*. However, in examples such as *Sam squeezed the rubber ball inside the jar* and *Sam urged Bill outside of the house* "neither the verbs *squeeze* or *urge* nor the prepositions *inside* or *outside* independently code motion" (Goldberg 1995: 158; italics original). In such cases, the construction coerces the "locative term into a directional reading" (Goldberg 1995: 159).

According to Michaelis, construction grammar is ideally suited to representing these coercion effects, since, unlike in modular models, special operators do not have to be employed. She notes that construction grammar "uses a single combinatory mechanism, the construction, to account for both coerced and syntactically transparent interpretations" (Michaelis 2003: 266). Furthermore, head-driven and template-based coercion are given a unified interpretation in the constructional model. In both cases, the construction's conventional meaning overrides the meanings associated with its components when they occur outside of the construction.

This begs the question then, how do we know which items can be successfully coerced into the construction and which will result in ungrammaticality? If the use of a particular item has not yet become conventionalized, then to a large extent its acceptability will vary from speaker to speaker. Nevertheless, as Goldberg (1995: 159) comments, "In order for coercion to be possible, there needs to be a relationship between the inherent meaning of the lexical items and the coerced interpretation". For example,

in the case of the caused motion construction, above, the relationship between the prepositional meaning and the constructional meaning is straightforward; the location given by the preposition is interpreted as the "endpoint of a path to that location" (Goldberg 1995: 159).

In this thesis, mismatch and coercion effects are shown to play an important role in the *it*-cleft's historical development. Over time, new subtypes of *it*-cleft have emerged (by extension from the prototype) in which more general patterns of correspondence are overridden. The coerced interpretations found in these mismatch sentences results in several new discourse functions for the *it*-cleft. As these new uses are conventionalized, the overarching *it*-cleft construction becomes more schematic and productive.

2.5 Noncompositionality, conventionality and construction-specificity

Noncompositionality is often viewed as the primary argument in favour of recognizing constructions as independent syntactic objects. If aspects of a construction's meaning cannot be broken down and attributed to its individual components, it cannot be said that constructions are epiphenomenal and they must instead gain theoretical significance. In construction grammar, symbolic relations are internal to the construction, linking form and meaning units not only at a componential level but also at a constructional level; that is, the complete structure is linked with all aspects of the construction's conventional meaning. The symbolic nature of the construction therefore allows it to have meaning which is not provided by or attributed to its individual components.

However, noncompositionality is often overestimated as a fundamental requirement for identifying a linguistic pattern as an independent construction. As Croft and Cruse (2004: 253) comment, "the common perception that a particular construction must be represented as an independent syntactic unit because it is 'noncompositional' is technically incorrect". In their work on idioms, Nunberg *et al.* (1994) suggest that many idiomatic constructions are in fact compositional. Using the example of *spill the beans*, the authors notes that this expression carries the meaning *divulge the information*. The construction's meaning is compositional in the sense that it can be separated and attributed to different constructional units. The syntactic element *spill* and the semantic

component *divulge* form a separate unit, while the noun phrase *the beans* is symbolically linked to the meaning *the information*. Despite this, Nunberg *et al.* (1994) argue that idiomatically combining expressions must still be identified as constructions, since these albeit compositional correspondences between form and meaning are nevertheless unique to this construction. As a result, the meaning of this expression cannot be determined from the general rules of semantic interpretation for these elements as they exist in other constructions.

Nunberg *et al.* (1994) therefore separate the concept of 'conventionality' from noncompositionality. For constructions to be identified as such, it is not necessary that their meaning cannot be broken down into components which can be attributed to individual formal elements, only that "their meaning or use can't be predicated, or at least entirely predicted, on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another" (Nunberg *et al.* 1994: 492).

Conventionality is therefore just as problematic as noncompositionality for componential theories of grammar. Regular syntactic expressions are both compositional and made up of form-meaning pairs that are general to other constructions. In a componential model, once we know the meanings that are typically associated with each element, we can correctly identify the meaning of these expressions. However, for less regular patterns, meaning may or may not be compositional, but crucially, the form-meaning pairings internal to the construction are specific to it. In a componential model, the constructional meaning of these patterns cannot be predicted from more general rules of the grammar. In a usage-based constructional approach, on the other hand, general as well as less regular patterns are captured by the constructional taxonomy, with the former existing at higher levels and the latter at lower levels in the inheritance hierarchy.

In this thesis, I claim that the *it*-cleft is a largely compositional construction. However, it is unpredictable to the extent that it contains construction-specific mappings between form and meaning. In this way, my analysis differs from the constructional account provided by Lambrecht (2001). He argues that the *it*-cleft requires a constructional analysis because of its noncompositional structure (see §6.2.1). In his

analysis, the *it*-cleft involves a complexity mismatch whereby the majority of syntactic elements are semantically empty and pragmatic meaning is constructional rather than compositional. Little attention is given as to how the *it*-cleft relates to other superordinate constructions in this account.

The approach taken in this thesis is therefore more in keeping with current models of how constructions work. Although construction-specificity and noncompositional structures are permitted, inheritance must also be prioritized. I suggest that the *it*-cleft is a motivated construction, inheriting from a larger specificational copular construction, which is in itself a mismatch structure. Furthermore, as I go on to explain in chapter 7, many of the *it*-cleft's construction-specific properties can be shown to have been originally inherited from constructions existing at earlier periods of the language. I liken this to the development of idioms such as *kith and kin*. The componential meanings for this idiom (*friends and family*) are now peculiar to this construction, but were once productive in the language. Likewise, in the *the X-er the Y-er* construction (such as *the more you practice, the easier it will get*), *the* is not a definite article, but comes from the Old English instrumental demonstrative *þy* (Fillmore, Kay and O'Connor 1988). According to Goldberg (1995: 119), in construction grammar, "Exceptions are allowed to exist, but only at a cost to the overall system". The account given in this thesis acknowledges this principle, and, as I argue, is all the better for it.

2.6 A note for the reader

As I explained in §2.1, the construction grammar framework I adopt in this thesis is consistent with the theories put forward by Croft (2001), Goldberg (1995, 2006) and Lakoff (1987). Like these authors, I use pictorial representations to elucidate my analyses. These diagrams have no formal status, serving only an illustrative purpose. Although some theories of construction grammar (in particular, Fillmore and Kay's Unification Construction Grammar) employ formal representations in the form of attribute-value matrixes (AVMs), I do not make use of this formalism. My pictorial representations hopefully make the analysis clearer, but there is no part of the analysis that cannot be stated in English prose and so there is no need for any given formalism.

3. TOWARDS A UNIFIED ACCOUNT OF SPECIFICATIONAL SENTENCES

This chapter lays the foundations for the analysis of *it*-clefts presented in chapter 4. Chapter 4 cannot be read in isolation from this precursory chapter, since together they amount to a unified analysis of specificational copular sentences. Throughout chapters 3 and 4, I build up an interwoven storyline culminating in three mutually supportive claims:

- First, I argue that *it*-clefts are a subtype of specificational sentence and that a unified account of specificational copular constructions can provide a maximally explanatory analysis of the *it*-cleft.
- Chapter 3 lays the groundwork for a unified analysis, focusing on the *NP be NP* specificational construction. I argue that specificational meaning results from a nominal predication relation which relies on the semantics of definite noun phrases.
- In chapter 4, I incorporate cleft sentences into this analysis. Applying my semantic account of specificational meaning to the *it*-cleft construction, it follows that *it*-clefts should also involve nominal predication. This leads me to argue for a (non-derivational) extraposition analysis of *it*-clefts in which the sentence-final clause is understood to be a restrictive relative, modifying the constituent *it*. Together, these items form a 'discontinuous constituent' which provides the definite-like description that is crucial to the interpretation of this structure as a specificational copular construction. I conclude that this analysis of *it*-clefts maximizes motivation from the language system (via inheritance from the more basic specificational schema) and has a number of advantages over alternative accounts.

Chapter 3 is structured as follows. In §3.1, I review a selection of *it*-cleft analyses that are labelled here, and elsewhere in the literature, as 'expletive' accounts. As I noted in §1.2.1, expletive analyses tend to maximize correspondence between *it*-clefts and noncopular subject-predicate sentences. In this section, I argue that such an approach is counterintuitive and leaves a number of questions regarding the structure

and function of this construction unresolved. I conclude that *it*-clefts should instead be examined in relation to other specificational copular constructions. Expletive analyses are characterized by their shared assumption that the postcopular element in *it*-clefts functions as the preposed argument of the proposition expressed in the sentence-final clause. As I go on to explain, a unified account of specificational copular sentences is not possible if we adopt an expletive analysis of *it*-clefts.

In §3.2, I claim that a satisfactory analysis of specificational sentences has not yet been provided. After reviewing both 'equative' and 'inverse' approaches and showing that neither can account for the full range of data, I argue that a better understanding of specificational meaning can help us to explain the behaviour of specificational copular constructions. However, as I go on to show, previous attempts at characterizing the semantic or pragmatic concept of 'specification' are not particularly useful for explaining both what is unique to these kinds of copular sentence and how they relate to other types of copular construction.

I provide my own original analysis of specificational copular constructions in §3.3. Working from the observation that definite noun phrases are frequently found in specificational sentences of the type *NP be NP*, I ask whether there is anything special about the semantics of definite descriptions. This leads me to argue that specificational sentences involve a special type of nominal predication relation. Unlike 'inverse' accounts, which also argue that specificational copular sentences involve predication, my constructional account does not invoke movement. As I explain, the analysis proposed in this thesis is able to both accommodate and explain a wider range of data than movement-based analyses.

3.1 Why examine *it*-clefts in relation to specificational structures?

As I noted in §1.1, there are four main elements in the syntactic structure of an *it*-cleft: the initial *it*, a form of the copular verb, a postcopular XP and a sentence-final clause. However, our understanding of how these components function and how they relate to one another differs from one analysis to another. As a result, it is common in the cleft literature for these structural subparts to be labelled using construction-specific, theory

neutral terminology. Following Hedberg (1990), I make use of the terms shown in (1) throughout this thesis.

(1) *Cleft pronoun* + *copula* + *clefted constituent* + *cleft clause*
It was Howard that left

In §1.2, I outlined the two main types of analysis that have been proposed to account for the *it*-cleft's structure: the expletive approach and the extraposition approach. In this section, I review some of the most influential expletive accounts in more detail, explaining why I choose to reject this type of analysis.

Expletive accounts maximize correspondence between *it*-clefts and their noncopular paraphrases. Common to all expletive accounts is the assumption that the cleft clause is directly predicated of (or is in some other way related to) the clefted constituent. This accounts for the truth-conditional equivalence between *it*-clefts and simple subject-predicate sentences. From this, it follows that the cleft pronoun and (in most accounts) the copular verb do not play an important role in the interpretation of the sentence.

(2) It was **Howard** that **left** [*it*-cleft]
(3) Howard left [canonical counterpart]

However, while all expletive accounts adopt the same set of basic assumptions, the details of these analyses differ. In particular, proponents of the expletive approach differ as to how they perceive the exact nature of the relationship between the postcopular element and the cleft clause and how they think this relationship comes about.

For Jespersen (1937: 86), *it*-clefts are syntactically identical to their canonical counterparts except for the addition of a "lesser subject and verb" and a "connective word", as shown in example (4), repeated from §1.2.1.[1]

(4) [It was] Howard [that] left
[sv] S [3^c] V

This "cleaving" or splitting up of the sentence by the semantically empty elements *it*, *be* and *that* serves only to highlight the clefted constituent and bring it into focus; it does not otherwise affect the interpretation or structure of the sentence proper. Consequently, the only real difference between the examples given in (2) and (3) above lies in their information structure: the elements *Howard* and *left* enter into the same type of predication relationship in both sentences.

Within the generative tradition, it is possible to express this predication relationship at different levels of representation. For example, Rochemont (1986) suggests that the clefted constituent is situated within the cleft clause at deep structure. At surface structure, this element is moved through the complementizer position within the cleft clause and out into the syntactic focus position adjoined to the verb *be*.

(5) $[_S$ It $[_{VP}$ $[_{V'}$ $[_V$ was$]$ $[_{NP}$ Howard$]_i]$ $[_{S'}$ t_i that $[_S$ t_i left$]]]]$

In other expletive accounts, the clefted constituent does not undergo movement, and is instead base-generated in the postcopular position. For example, Chomsky (1977) proposes that the clefted XP is base-generated in a topic position adjoined to the clause formed on S'. This constituent is coindexed with the gap, generated by *wh*-movement, in the sentence-final clause, shown in (6).

(6) $[_S$ It $[_{VP}$ was $[_{S''}$ $[_{TOP}$ $[_{NP}$ Howard$]]_i$ $[_{S'}$ wh_i that $[t_i$ left$]]]]]$

[1] The use of square brackets in Jespersen's (1937: 86) formalism indicates that this information is "extraposed" relative to the sentence proper. Lower case s and v indicate the "lesser" (i.e. expletive) subject and verb and 3^c represents a "tertiary connective".

For Chomsky then, the *it*-cleft is a type of topicalization construction, involving *wh*-movement. The relationship between the clefted constituent and the cleft clause is therefore akin that between topic and comment.

Williams (1980) argues that the clefted constituent and the cleft clause enter into a subject-predicate relation which is indicated by indexing. He claims that the clefted constituent and the cleft clause are coindexed at surface structure as a result of the "grammatically governed environment of predication" in which X (any maximal projection) precedes an S' configuration. For Williams (1980: 229), *it*-clefts are "the mirror image of extraposed sentences". While in the *it*-cleft in (7), the clefted element is the subject and the clausal constituent is the predicate, in the extraposed sentence in (8), the postcopular element is the predicate and the sentence-final clause is the subject. In both structures, the constituent *it* is analysed as an "inert element which does not count as subject" (Williams 1980: 221).

(7) It was Howard that left

Subject Predicate

(8) It is clear that he left

Predicate Subject

According to Delahunty (1982, 1984), on the other hand, the predication relation between the postcopular element and the cleft clause takes place solely at the level of logical form. He assumes that because *it*-clefts and their noncopular subject-predicate counterparts are truth-conditionally synonymous, they must have equivalent Logical Structures. He achieves this by converting the cleft clause into a function. The free variable contained within this clause is bound by the lambda operator, creating an abstract: the function in (10). The Logical Structure given in (10) states that the value for the variable (x) (in *x left*) is denoted by substituting (h) (*Howard*) for x. By reducing this

lambda expression to its "normal form", we obtain the predication relation of the simple sentence *Howard left*, given in (11).

(9) It was Howard that left

(10) (h) [λx[L(x)]]

(11) L(h)

However it achieved, whether underlyingly at deep structure (Rochemont 1986), via indexing at surface structure (Chomsky 1977; Williams 1980) or as a result of lambda conversion at the level of logical form (Delahunty 1982, 1984), all of the above analyses have in common the notion that the cleft clause is directly predicated of or is in some way related to the postcopular XP[2]. In this way, expletive accounts highlight the similarities and maximize the correspondences between *it*-clefts and their **noncopular** subject-predicate counterparts[3]. In other words, although these authors may recognize that *it*-clefts are subspecies of **copular** sentence, this fact does not form the basis of their analysis.

An exception to this generalization is Heggie (1988) who integrates her expletive account of *it*-clefts into a general theory of copular sentences. As I explained in §1.1, *it*-clefts have a specificational, or identifying, meaning and can therefore be regarded as a subtype of specificational copular sentence. Adopting an 'inverse' analysis of specificational structures, Heggie assumes that specificational copular sentences, such as (13), are the opposite of predicational copular sentences, such as (12); while in predicational sentences, the predicate occurs to the right of *be*, in specificational sentences, the precopular phrase is the predicate (see §3.2.2 for further details on the inverse account).

---

[2] In É. Kiss' (1998) expletive analysis, the clefted constituent can be either moved into the postcopular position from within the cleft clause or base-generated and linked to the corresponding *wh*-pronoun in the cleft clause at LF. The purpose of this dual derivation is that it enables É. Kiss to accommodate the *it*-cleft's unusual agreement patterns. I discuss this aspect of É. Kiss' (1998) analysis in §4.3.

[3] However, as I explained above, for Chomsky (1977), the clefted constituent and the cleft clause enter into a topic-comment, rather than subject-predicate, relationship.

(12) John is evil/a murderer [predicational copular sentence]

(13) The murderer is John [specificational copular sentence]

From this, it follows that specificational copular sentences always involve a predication relationship. Since most expletive accounts assume that the cleft clause is directly predicated of the clefted constituent, Heggie argues that *it*-clefts can therefore be integrated into a unified analysis of copular sentences.

Heggie claims that the copular verb (in all copular sentences) is a raising verb which takes a small clause complement. The copula functions as a "verbal operator" which creates a predicate out of any phrasal category via coindexing. This index then "spreads" to the subject of the small clause via predication (following Williams 1980). In the case of the *it*-cleft, the sentence-final clause becomes the predicate and the postcopular focal element functions as the subject of the small clause. As Heggie (1988: 183) comments, this verbal operator function of *be* is so unrestricted that it "allows *be* to create small clause structures which do not exist in any other context", such as the CP-small clause found in clefts.

Despite their differences, these expletive accounts all share the same fundamental problems. As I explained above, expletive analyses are so-called because they assume that the cleft pronoun is a semantically empty dummy element. However, this conclusion is not supported by the cleft data. Aside from the initial *it*, the cleft structure can be introduced by other elements. Crucially, this exchange results in a difference in meaning. For example, the initial *it* in *it*-clefts can alternate with the demonstrative pronouns *this* and *that* (Hedberg 2000). These demonstrative clefts are used to indicate temporal or spatial deixis within the immediate discourse context, shown in (14) and (15).[4]

(14) This is Sarah we're talking about, isn't it?

[4] I use the term 'demonstrative cleft' here, and throughout this thesis, to refer to examples which exhibit the same configuration as the *it*-cleft, but which are introduced by demonstrative pronouns. This term is used in a difference sense in Calude (2007, 2008) to refer to a subset of reverse pseudoclefts with demonstrative subjects, such as *that's what I thought*.

(15) That was Tom who just phoned

In addition to *it*-clefts and demonstrative clefts, we also have *there*-clefts. Unlike *it*-clefts, *there*-clefts do not carry an exhaustiveness implicature. For instance, we assume from the *it*-cleft in (16) that *Sally* is the only person *working Saturday*. However, the *there*-cleft in (17) is used for the purpose of listing a potentially incomplete list of possible candidates.

(16) It's Sally that's working Saturday

(17) Well, there's Tom and Sally that are available to work Saturday

Cleft constructions containing elements other than *it* as the cleft pronoun therefore result in a difference in meaning and/or discourse function. This data therefore suggests that the initial *it* of the *it*-cleft is not a meaningless dummy element and in fact plays an important role in the interpretation of the sentence. For example, the lack of an exhaustiveness implicature in *there*-clefts suggests that the property of exhaustiveness in *it*-clefts may be attributed to the constituent *it*. Consequently, by completely ignoring the contribution of *it* and focusing only on the relationship between the clefted constituent and the cleft clause, an expletive analysis of *it*-clefts is really an incomplete analysis.

In addition, it troubles me that although the cleft clause looks and behaves like a restrictive relative clause, expletive accounts assume that this is really a non-modifying sentential predicate. For Delahunty (1982), Rochemont (1986) and Heggie (1988), relative pronouns, like *who*, should only occur in the cleft clause as a result of analogy with relative clauses. Chomsky (1977) and Williams (1980), on the other hand, claim that the cleft clause has the same internal structure as restrictive relatives. This is supported by the fact that they are subject to the same general constraints on *wh*-movement (see Chomsky 1977).[5] Nevertheless, since the cleft clause does not

[5] For example, as shown in (i) and (ii), both structures obey the complex-NP constraint (Chomsky 1977).

(i) *It's this document that I accept the recommendation that I should sign
(ii) *The document that I accept the recommendation that I should sign is over here

restrictively modify its antecedent, it is said to provide a different function from other restrictive relatives.

However, the expletive approach seems to be missing an important generalization here. Rather than trying to explain away or limit the similarities between the cleft clause and restrictive relatives, we should instead be trying to apply what we know about restrictive relative clauses to *it*-clefts. If the cleft clause is structured internally like a restrictive relative, then the chances are that it also has a modifying function. If this clause does not modify the postcopular constituent then we have to ask, what does it modify?

Furthermore, as syntax-centred approaches, the expletive analyses outlined above do not focus on explaining how the *it*-cleft's pragmatic properties, such as exhaustiveness and existentiality, come about. An exception to this generalization is the expletive account provided by É. Kiss (1998). She claims that "exhaustive identification is a function of structural focus" (É. Kiss 1998: 251). Identificational focus (which expresses exhaustive identification) has a designated syntactic position, occupying the specifier slot of the focus phrase. In É. Kiss' analysis, the clefted constituent occupies this scope position, thereby performing two important roles.

(18) $[_{IP}$ It $[_{I'}$ was$_k$ $[_{FP}$ HOWARD$_i$ $[_{F'}$ t_k $[_{CP}$ $[t_i$ $[_{C'}$ that $[_{IP}$ t_i left]]]]]]]]

Syntactically, the clefted constituent functions as an operator which "marks the sentence part following it and c-commanded by it as the scope of exhaustive identification" (É. Kiss 1998: 253). Semantically, the clefted XP expresses the complete set of elements of which the predicate in the cleft clause holds. In this way, É. Kiss provides a syntactic explanation for the fact that *it*-clefts carry exhaustiveness implicatures.

I review É. Kiss' (1998) analysis in §4.1 amid a larger discussion which compares alternative accounts of the exhaustiveness in *it*-clefts. After identifying a number of fundamental problems with É. Kiss' syntax-based analysis, I argue that exhaustiveness in *it*-clefts is in fact a semantic phenomenon and is the outcome of universal quantification. Since in all expletive accounts, the initial *it* is semantically

empty, no remaining constituents exist which could be said to perform this quantifying function. As a result, the expletive approach does not provide us with a satisfactory explanation for how the *it*-cleft's meaning differs from that of a simple, noncopular, subject-predicate sentence.

However, as I see it, the most important problem with the expletive approach is that it does not prioritize the relationship between *it*-clefts and other specificational copular sentences. I explained above that Heggie (1988: 184) considers her account to be the exception. She suggests that her expletive account can be integrated into a unified analysis of copular sentences. Assuming an 'inverse' analysis of specificational sentences, Heggie claims that all copular constructions involve predication. However, there is an important problem with Heggie's story. In the inverse account, specificational copular sentences always involve nominal predication. For example, in (13), repeated here as (19), the precopular NP *the murderer* is the predicate and the postcopular element *John* is the argument.

(19) The murderer is John

Since *it*-clefts are a subtype of specificational copular sentence, we would assume that these too involve nominal predication. In Heggie's expletive analysis, on the other hand, cleft sentences contain a CP-small clause. Predication of this sort is not shared by other specificational copular constructions and this property of the *it*-cleft is therefore an unexplained exception to a generalization.

Despite Heggie's (1988) attempt, expletive analyses do not allow for a truly unified account of specificational copular sentences. Simply stating that all kinds of specificational copular sentence involve one type of predication or another is not sufficient, since this property is shared by numerous other non-specificational and noncopular sentence types. In order to provide a unified analysis of specificational sentences we need much more than this. Throughout the rest of this chapter, and in the next, I build up my own unified analysis of specificational copular sentences (including *it*-clefts). My starting point is to create a more accurate understanding of what

specificational meaning in copular constructions actually is. This allows me to be more precise about the basic components that every specificational copular construction must contain and how this family of constructions differs from noncopular as well as non-specificational sentence types. In chapter 4, I show that *it*-clefts are best treated as a subtype of specificational sentence and that this approach explains a number of otherwise obscure facts about the construction. The upshot is a (non-derivational) extraposition account of *it*-clefts. As I go on to explain, this analysis does not suffer from the same shortcomings as the expletive accounts outlined above.

3.2 What is specificational meaning?

In §3.3, I present my own, largely semantic, account of *NP be NP* specificational copular sentences. As a precursor to this, in this section, I outline some of the different kinds of analysis that have been proposed in the literature. I begin by presenting two opposing analyses of specificational structures: the 'equative' analysis and the 'inverse' analysis. While an equative analysis assumes that two referring expressions are equated in a specificational sentence, proponents of the inverse analysis instead suggest that these sentences involve a predication relationship. I then go on to discuss two less formal accounts of specificational meaning proposed by Higgins (1979) and Declerck (1988). From this, I conclude that the concept of specificational meaning is not well understood and that the semantic contribution of the components of specificational sentences has yet to be given an adequate description.

As I noted in chapter 1, specificational copular sentences include noncleft examples as well as pseudoclefts (including *wh*-clefts and *th*-clefts) and *it*-clefts. These constructions all have a similar 'identifying' function. For instance, all of the examples below identify *John* as *the murderer*.

(20) The murderer is John

(21) The one that murdered Sally was John

(22) It was John that murdered Sally

Throughout the literature, it is agreed that specificational copular sentences have a different meaning from predicational copular sentences. For example, in (23), the predicative noun phrase *a murderer* 'describes' the referent *John*, or 'ascribes' a property to *John*.

(23) John is a murderer

However, what is not agreed upon is how specificational (or identifying) meaning comes about. To answer this question, we need to know what the semantic contribution is of the individual components of a specificational sentence.

3.2.1 The equative analysis

For some authors, specificational copular sentences are treated in the same way as equatives. Equative sentences (or identity statements) contain two phrasal constituents of the same semantic type which are involved in a relationship of identity. For example, in (24), both *Cicero* and *Tully* are referring expressions; the proposition states that the person we know as *Cicero* and the person we know as *Tully* are actually one and the same individual. Authors differ as to how this relationship of identity comes about. For example, while some suggest that equative sentences involve a special "equative" or "identifying *be*", others argue that the copular verb is always semantically inert and that the identity relationship instead originates from a special "equative small clause" structure (see Heycock and Kroch 1999). Either way, by extending the equative analysis to specificational sentences, both *the murderer* and *John* in (25) are treated as referring expressions.

(24) Cicero is Tully

(25) The murderer is John

The equative analysis of specificational copular sentences has intuitive appeal since an identity relationship would explain why specificational sentences have an

identifying meaning. In addition, both equatives and specificational sentences are reversible. For example, (24) and (25) above have the same meaning as (26) and (27), respectively. In (26), the individuals *Tully* and *Cicero* are identified as one and the same person and in (27), the focal element *John* is identified as *the murderer*. For proponents of the equative analysis of specificational sentences, predicational copular sentences do not share this property. For example, the subject and predicate of (23) above cannot be reversed without leading to ungrammaticality, shown in (28).

(26) Tully is Cicero

(27) JOHN is the murderer

(28) *A murderer is John

Based upon this evidence, Heycock and Kroch (1999: 380) argue "that it is not possible to treat any constituent appearing in [the subject position of a copular sentence] as predicated of a postcopular argument". This claim represents a challenge to 'inverse' accounts (outlined below) which analyse the subject of a specificational sentence as an underlying predicate. In an equative analysis, the noun phrases of both equative and specificational sentences are reversible because they are of the same semantic type. As a result, it should not matter in which order they appear.

Heycock and Kroch (1999, 2002) are proponents of the equative analysis. They suggest that this approach has an "economy advantage", since it reduces the taxonomy of copular sentences to just two types: predicational and equative. However, a potential problem for the equative analysis is that the phrasal components of specificational sentences do not seem to be as "equal" as those of "true" equatives. Heycock and Kroch concede that one of the noun phrases in a specificational sentence is usually "less referential" than the other. For example, in (25) above, while the postcopular proper name *John* is clearly a referring expression, the definite noun phrase subject *the murderer* does not seem to exhibit this property, or at least not to the same extent. These authors note that the "assimilation of specificational sentences to equatives runs afoul of

the intuition that the former are asymmetric in interpretation in a way that "true equatives" are not" (Heycock and Kroch 1999: 381).

As a way of getting around this problem, Heycock and Kroch (1999, 2002) claim that the asymmetry in specificational sentences is not in fact semantic, but is the result of the focus-ground opposition, which causes each noun phrase to differ in informativeness. Consequently, these authors maintain that specificational sentences contain two constituents of the same semantic type. I discuss the equative approach in more detail in §5.2, where I compare it to my own analysis of specificational copular constructions.

3.2.2 The inverse analysis

An alternative account views specificational copular sentences as inverse predicational sentences of the type *NP be NP*. Proponents of this view include Williams (1983), Partee (1986), Heggie (1988), Moro (1997), Mikkelsen (2005) and Den Dikken (2006). In these analyses, the copular sentence in (29) is predicational, with the postcopular noun phrase *the best doctor* describing the referent *John*. However, for the corresponding specificational sentence in (30), the predicative element is in initial position. In contrast to the equative analysis then, the two noun phrases of a specificational sentence are involved in an asymmetric predication relationship and the initial definite noun phrase is non-referring.

(29) John is the best doctor

(30) The best doctor is John

In some versions of this analysis, both specificational and predicational sentences are derived from the same small clause structure, in which the referential NP always precedes the predicative noun phrase. For example, Moro (1997) and Mikkelsen (2005) propose that while predicational sentences are derived via the raising of the referential noun phrase into the subject position, the predicative noun phrase is raised to derive a

specificational sentence.[6] In this syntactic derivational account, the distinction between predicational (descriptive) sentences and specificational (identifying) examples is therefore dependent upon the alignment of the referential and predicative noun phrases.

One advantage to the inverse analysis is that it acknowledges, and provides an explanation for, the asymmetry between the two noun phrases of a specificational sentence. In addition, there is good evidence for the claim that specificational sentences contain non-referring subjects. By placing these sentence types in environments involving left-dislocation structures, question-answer pairs and tag questions, Mikkelsen (2002, 2005) shows that while in predicational and equative constructions, anaphoric pronouns show agreement with the subject in terms of gender, number and animacy, specificational sentences allow the pronoun *it* to be anaphoric to gender-specific subjects.

(31) The tallest girl in the class is Swedish, isn't **she**? [predicational]

(32) The tallest girl in the class is Molly, isn't **it**? [specificational]

(33) [Pointing to a player on the field]
SHE is Molly Jacobson, isn't **she**? [equative]

(examples from Mikkelsen 2005: 72)

These results indicate that there is a definite semantic difference between the subjects of specificational and equative sentences. As Heycock and Kroch (2002: 106) acknowledge, "This behaviour is quite unexpected under an equative analysis". On the basis of her pronominalization evidence, Mikkelsen claims that there are three distinct types of copular sentence: predicational (involving a referential subject and a predicate), specificational (in which the predicative element is raised into the subject position and the referring expression is located in the postcopular position), and equative (containing two referring expressions).

[6] Adger and Ramchand (2003) make use of a similar analysis in order to account for an inverted predication construction found in Scottish Gaelic. In support of the inverse analysis, they conclude that all predication constructions reduce to a single underlying syntactic structure.

For Mikkelsen (2002, 2005) then, specificational and equative copular sentences cannot be given a unified analysis; while specificational sentences are inverse predicational structures, the semantically equivalent phrases in equative sentences are not involved in a predication relationship. However, this tripartite taxonomy of copular sentences is not advocated in all versions of the inverse approach. For proponents of the inverse analysis who are interested in creating a maximally economical account of copular sentences, equatives are also analysed as predicational structures (see for example, Partee 1986, Heggie 1988 and Den Dikken 2006). I discuss these so-called "economical" accounts in more detail in §5.2.

Despite its intuitive appeal and the support of pronominalization evidence, there are a number of problems with the inverse analysis of specificational sentences. A common objection is that not all NP predicates can occur in subject position. For example, as I noted above, predicational copular sentences with indefinite noun phrase predicates, such as (34), typically resist "inversion". As I go on to show in §3.3.4, proponents of the inverse analysis have failed to provide an adequate description of, let alone an explanation for, these restrictions.

(34) John is a murderer

(35) *A murderer is John

More importantly for my concerns, the inverse analysis does not solve, or even broach, the question of how specificational meaning comes about. The inverse analysis assumes that specificational sentences are derived from the same structure as predicational copular sentences. From this, we must conclude that specificational meaning is the result of a syntactic movement operation. The raising of the referential element into subject position leads to a predicational copular sentence, which tells us "something **about** the referent of the subject", while the raising of the predicative element results in a specificational sentence, telling us "**who** or **what** the referent is" (Mikkelsen 2005: 1; emphasis original). However, it is not clear exactly how or why this syntactic operation would result in this well noted difference in meaning. If

specificational copular sentences are inverse predicational structures, why don't they also describe, rather than identify, the referent?

Furthermore, if we assume that specificational meaning is dependent upon a particular inverse word order, it follows that once we reverse the noun phrases of a specificational sentence the result should automatically be predicational. For instance, as shown in the examples below, Mikkelsen (2005: 177) categorizes all asymmetric copular sentences introduced by referential subjects as predicational.

(36) A. Who is the winner? B. JOHN is the winner, isn't he? [predicational]

(37) A. Who is the winner? B. The winner is John, isn't it? [specificational]

(examples from Mikkelsen 2005: 177)

However, as noted by Declerck (1988: 93), examples such as (36), which contain a focal referring subject, have the same specificational function of identifying rather than describing the referent as the "inverse" structure in (37). Likewise, Partee (1986: 363) acknowledges that her inverse analysis leaves open the question of how to deal with examples such as (38), which have an "unambiguously" specificational rather than predicational meaning.

(38) Unusual is what John is (Partee 1986: 363)

This begs the question then, how can we explain these "reverse specificational" sentences in a syntactic movement account such as the inverse analysis?

3.2.3 Two less formal accounts of specificational meaning

Both the equative and the inverse analyses of specificational copular sentences therefore overlook important semantic differences between, on the one hand, equative and specificational sentences, and on the other, specificational and predicational sentences. In order to create a unified analysis, equative accounts "work around" the obvious asymmetry in specificational examples and fail to account for the pronominalization

evidence which suggests that the subjects of specificational sentences are non-referring. In the syntax-based inverse account, the identifying function of specificational sentences is not predicted and consequently, no explanation is provided as to how this meaning arises.

However, two less formal accounts of specificational sentences have been proposed which focus on characterizing the nature of specificational meaning. In contrast to the equative and inverse analyses, Higgins (1979) and Declerck (1988) argue that specificational copular sentences involve neither identity nor predication relationships. Since these authors are concerned with providing a detailed description of copular sentences rather than a unified and economical analysis, both Higgins (1978) and Declerck (1988) propose taxonomies of copular sentences that go beyond the predicational, specificational and equative types that we have considered so far. For example, Higgins (1979) recognizes an additional 'identificational' type of copular sentence and Declerck (1988) suggests that we need to distinguish 'descriptionally-identifying' and 'definition' types as well. However, these are not well-defined, uniform categories and many of the examples they include can be accommodated into the tripartite taxonomy of predicational, equative and specificational sentences.[7] Consequently, I do not discuss these alternative types of copular sentence in any more detail here.

3.2.3.1 Higgins (1979)

Higgins (1979) argues that specificational sentences function like lists; that is, the subject of a specificational sentence acts as the heading of the list and the postcopular elements serve as items on that list. For instance, the noncleft in (39) and the *wh*-cleft in (41) can, according to Higgins, be paraphrased as the lists given in (40) and (42), respectively.

---

[7] For example, many of Declerck's (1988) 'descriptionally-identifying' tokens can be classified as predicational copular sentences. Example (i), for instance, describes (or classifies) the referent (*that man*) as *John's brother*.

(i) That man is John's Brother (Declerck 1988: 95)

(39) John's opponents were Sally Simons and Reese Winters

(40) John's opponents: Sally Simons, Reese Winters

(41) What I don't like about John is his tie (Higgins 1979: 214)

(42) Things I don't like about John: his tie

Higgins suggests that specificational sentences involve a 'value-variable' relation. He notes, "the heading of a list provides a "variable", thereby delimiting a certain domain, to which the items on the list conform as "values" of that variable" (Higgins 1979: 155). Higgins labels the precopular noun phrase using his own term 'superscriptional' and refers to the postcopular element as 'specificational'. Superscriptional elements can be definite or indefinite noun phrases, while all types of constituent may be specificational.

Aside from arguing that specificational sentences function as lists, Higgins is not forthcoming when it comes to defining what characterizes the 'superscriptional' and 'specificational' components. However, by using his own terminology, Higgins makes it clear that neither the precopular nor the postcopular element of a specificational sentence can be characterized as 'referential' or 'predicational'. For example, Higgins separates specificational sentences from identity statements involving two referring expressions. He notes that specificational meaning "is not the expression of some kind of identity" (Higgins 1979: 214). Likewise, Higgins (1979: 214) argues that specificational sentences do not involve a predication relationship, since "The whole notion of being "about" something is alien to a list".

However, this does not get us any closer to finding out what 'superscriptional' noun phrases actually are and how they relate to more well-defined concepts such as 'referential' and 'predicational'. This must be achieved if we are to find an explanation for the connection between specificational and predicational copular sentences that Higgins draws our attention to. For example, he assumes that although the pair of sentences in (43) and (44) are not directly related, "the existence of the parallelism clearly means something. A correct characterization of Superscriptional noun phrases should account for this" (Higgins 1979: 274-275). However, because Higgins does not

have an accurate definition of the term 'superscriptional', he cannot provide an explanation for this phenomenon.

(43) That he hasn't come is a problem [predicational]

(44) The problem is that he hasn't come [specificational]

(examples from Higgins 1979: 274)

As a result of using ill-defined, construction-specific terminology, Higgins' account of specificational copular sentences is therefore of little practical value when it comes to drawing generalizations or asking how the different kinds of copular sentence relate to one another.

3.2.3.2 Declerck (1988)

An alternative account of specificational meaning is offered by Declerck (1988). Declerck agrees with Higgins (1979) that specificational sentences involve a 'value-variable' relationship. However, by extending this concept, Declerck provides an extremely broad definition of specificational meaning as pertaining to any sentence, copular or noncopular, that gives the answer to a *wh*-question. So, for example, both sentences in (45) are specificational for Declerck, providing the 'value' *John* for a 'variable'. The noncopular sentence in (45a) is an example of Declerck's 'type 1' specificational sentence. Here the variable is not present in the formulation of the sentence, but is described by the verb phrase *committed the murder*. The *th*-cleft given in (45b), on the other hand, is an example of a 'type 2' specificational sentence, with the subject noun phrase containing the variable that the value specifies, in this case *the x who committed the murder*. Unlike type 1 specificational sentences, type 2 examples are reversible, shown in (46).

(45) Who committed the murder?

a) JOHN committed the murder (Declerck 1988: 8)

b) The one who committed the murder is JOHN

(46) JOHN is the one who committed the murder

For Declerck then, specificational meaning is not peculiar to copular sentences. More importantly however, even within the taxonomy of copular sentences, specificational meaning is not limited to the specificational copular construction. Declerck argues that predicational copular sentences can also have a specificational meaning as long as they answer a *wh*-question. For example, although (47) is a predicational copular sentence, it is also a type 1 specificational sentence having predicational focus. Predicational copular sentences therefore make up a distinct construction from specificational copular sentences, not as a result of their distinct semantics, but because only the latter, as type 2 specificational sentences, are reversible.

(47) What is John like?
John is SILLY (Declerck 1988: 39)

Declerck therefore reduces the concept of specificational meaning to contrastive focus. Consequently, in this account, the identifying function of specificational copular sentences is purely information-structural and is not the product of the semantic contribution of its components. As a result, the definition of specificational meaning as a 'value-variable' relationship cannot help us to identify what distinguishes specificational (identifying) copular sentences from predicational (descriptive) copular sentences. For example, although (47) above has contrastive focus, the predicative element *silly* nevertheless 'describes' *John*. If we want to use specificational meaning as means of characterizing the components of a specificational copular sentence and to help us to understand the structural properties of these sentences, we need a semantic account of specificational meaning that is specific to this construction and yet, *contra* Higgins (1979), is described in terms of general, well-understood concepts.

Declerck's (1988) characterization of the components of a specificational copular sentence is also problematic. He suggests that specificational copular sentences contain one "weakly referring" noun phrase and one "strongly referring" element. As a result,

although Declerck (1998: 3) argues that such sentences do not "state a relation of identity between two entities" and cannot therefore be conflated with identity statements, his analysis is nevertheless very similar to an equative account, which analyses specificational sentences as containing two referring expressions.

Declerck relies on Donnellan's (1966) referential/attributive distinction to capture the asymmetry between the two referring expressions of a specificational sentence. Donnellan notes that in an example such as (48), the initial definite description has two possible interpretations. If the speaker knows that for example *John* is *Smith's murderer*, then this definite description is treated as 'referential' by Donnellan. If, however, the speaker does not know *Smith's murderer*, they may still utter (48) to indicate that whoever the murderer is, they are insane. In such cases, Donnellan labels the use of the definite description in subject position as 'attributive'. As Donnellan (1966: 285) puts it, "A speaker who uses a definite description attributively in an assertion states something about whoever or whatever is the so-and-so".

(48) Smith's murderer is insane (Donnellan 1966: 285)

Declerck (1988) claims that the 'variable' in specificational copular sentences is a weakly referring 'attributive' noun phrase, while the 'value' is a strongly referring expression. He suggests that "The purpose of a specificational sentence with a definite NP as a variable is precisely to identify the referent of an NP whose use has remained attributive for the hearer" (Declerck 1988: 47). However, Declerck recognizes that an attributive noun phrase cannot distinguish specificational copular sentences from other copular constructions. For example, he argues against Elffers (1979), who claims that an attributive subject noun phrase is "a sufficient condition for the sentence to be specificational", since predicational copular sentences, such as (48) above, can also contain attributive subjects (Declerck 1988: 48f).

However, this fact presents a problem for Declerck (1988), since it means that his analysis cannot explain Mikkelsen's (2005) pronominalization evidence. As I explained in §3.2.2, these results suggest that the subjects of specificational sentences

are semantically distinct from the subjects of predicational and equative sentences. Mikkelsen (2005: 89) shows that in predicational copular sentences even subjects with "unknown reference" (our 'attributive' noun phrases) "do not pronominalize as *it*, but with the gendered pronouns *she/he*". For example, gendered pronouns are used to refer back to *Smith's murderer* in (49) regardless of whether the speech participants know the identity of this person or not. In contrast, the subjects of specificational sentences invariably pronominalize with *it*, as in (50).

(49) Smith's murderer (whoever he is) is insane, isn't he?

(50) Smith's murderer is John, isn't it?

This evidence suggests that the subject of a specificational sentence is neither strongly nor weakly referring and consequently cannot be satisfactory categorized as an attributive noun phrase. Declerck's account therefore suffers from the same problems as other equative analyses.

3.2.4 Interim summary

In conclusion then, this short review of some of the different kinds of approaches to specificational copular sentences has shown that none of the existing analyses have managed to provide an account that tells us **both** how specificational subjects differ from the subjects of other copular constructions **and** how specificational meaning results from the semantic contribution of the construction's components. In the following section, §3.3, I present my own, largely semantic account of specificational copular sentences. In contrast to Declerck's (1988) all-encompassing 'value-variable' relationship, I argue for a construction-specific definition of specificational meaning as resulting from a special type of nominal predication relationship.

By assuming that specificational sentences involve nominal predication, my analysis has much in common with the inverse analyses outlined in §3.2.2. However, unlike inverse accounts, my analysis does not rely on syntactic movement. As I go on to explain, the semantic account offered in this thesis therefore profits from the insights of

the inverse analysis without encountering the same problems. I frame my discussion in response to Mikkelsen (2005). This choice is based on a number of reasons. Mikkelsen (2005) is a recent example of the inverse approach which builds upon earlier accounts, such as Moro (1997). Unlike Den Dikken (2006), who develops his own 'relators and linkers' approach to predication, Mikkelsen (2005) is therefore representative of the "standard" inverse analysis. In addition, I agree with Mikkelsen that there are three distinct types of copular sentence: predicational, specificational and equative. I argue in §5.2 that attempts to subsume equative and specificational sentences within the same inverse predicational structure involve a false economy. By responding to Mikkelsen (2005), I can therefore focus on the key issue of comparing a semantic predicational account to one involving syntactic movement.

3.3 My own analysis of specificational sentences

In this section, I develop a largely semantic, constructional account of specificational copular sentences of the type *NP be NP*.[8] My argument is structured as follows. In §3.3.1, I outline some of the difficult data that, to my mind, a satisfactory account must be able to accommodate and/or explain. With this goal in place, I suggest that a profitable starting point is to examine the semantic role of definite descriptions in specificational sentences. In §3.3.2, I provide a Russellian analysis of definite noun phrases, before going on, in §3.3.3, to incorporate this into a semantic account of specificational constructions. I conclude that specificational meaning results from a nominal predication relation involving definite noun phrases. However, while this analysis therefore retains the intuition of inverse accounts that specificational sentences are semantically asymmetrical, it does not rely on syntactic movement. As I explain, the account provided in this thesis is therefore able to accommodate more data than movement-based analyses, since it predicts that specificational meaning is independent

[8] Noncleft specificational sentences do not all have the form *NP be NP*. Nevertheless, these sentences form the prototypical case and make for a convenient starting point for our analysis. In what follows, I argue that specificational sentences always contain a predicative NP. However, the other phrasal element does not always have to be nominal. Specificational sentences containing non-NP phrasal elements are not fully integrated into my account until chapter 8.

from word order. Finally, in §3.3.4, I provide an explanation for the limited distribution of indefinite noun phrase subjects based upon the differing semantics of definite and indefinite noun phrases. I argue that this semantic account explains the data better than the information-structural distinction that Mikkelsen (2005) invokes.[9]

3.3.1 The data

As I see it, there are three sets of data that a satisfactory analysis of specificational *NP be NP* sentences must be able to accommodate and/or explain. The first set of data is Mikkelsen's (2005: 72) pronominalization evidence, introduced in §3.2.2. This data suggests that specificational sentences differ from predicational copular sentences with respect to the semantic content of their initial noun phrase. For example, in the predicational sentence given here as (51), the pronoun *she* agrees with the referential subject. However, in (52), this same gender-specific definite noun phrase pronominalizes with *it*.

(51) The tallest girl in the class is Swedish, isn't **she**? [predicational]

(52) The tallest girl in the class is Molly, isn't **it**? [specificational]

(examples from Mikkelsen 2005: 72)

Since *it* appears to be the default pronoun for non-referring, predicative expressions (as shown in (53)) this evidence leads me to conclude that while the subject of the predicational sentence in (51) is semantically referential, the initial NP in (52) is a non-referring expression. On this basis, I reject the equative account, which analyses (52) as containing two referential noun phrases, and will assume that specificational sentences are semantically asymmetrical. The question therefore arises as to how the initial definite noun phrase in (52) obtains a non-referring semantics?

(53) Molly is very tall. **It** is a difficult thing to be at her age.

[9] The analysis outlined in §3.3 builds upon ideas originally presented as a conference paper entitled 'How specificational are cleft sentences?' (see Patten 2007).

As I noted in §3.2.2, Mikkelsen's (2005) inverse account treats specificational copular sentences as "reverse" predicational sentences. For Mikkelsen (2005), both specificational sentences and predicational copular sentences of the type *NP be NP* are derived from the same small clause structure (shown in Figure 3.1).

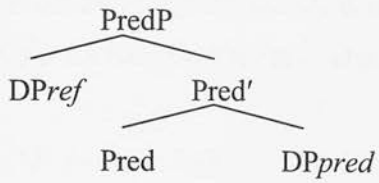


Figure 3.1 *The predicational core (Mikkelsen 2005: 166)*

The argument goes that if the referential noun phrase is raised into the subject position, a predicational copular sentence is obtained, while a specificational copular sentence results from a movement operation which raises the predicative noun phrase. So, in this analysis, the initial NP of the specificational copular sentence in (55) corresponds to the predicative noun phrase of the predicational sentence in (54).

(54) John is the best doctor, isn't he?

(55) The best doctor is John, isn't it?

This inverse analysis neatly explains Mikkelsen's pronominalization results. However, I have chosen to reject a movement-based account. As I explained in §3.2.2, the inverse account equates specificational meaning with an "inverse" word order, in which a predicate precedes its argument. But, this cannot explain reverse specificational sentences, such as (56), which contain referential subjects but still have a specificational (or identifying) meaning.

(56) Who is the best doctor? JOHN is the best doctor, isn't he?

In order to accommodate this data, we need to find an analysis of specificational *NP be NP* sentences that can explain how specificational meaning comes about without relying on syntactic movement or word order to do so. To my mind, the final set of data that is in need of an explanation can actually help us to achieve this goal.

In §3.2, I noted that specificational sentences with initial indefinite noun phrases are often deemed ungrammatical, as shown in (57). For inverse accounts, which assume that predicates can precede their subjects, this data is problematic.

(57) *A doctor is John

Mikkelsen (2005) attempts to explain this data by arguing that predicates can only move to subject position if they contain discourse-old information; since indefinite noun phrases are typically associated with discourse-new information, they rarely meet this criterion. However, in §3.3.4, I argue that Mikkelsen's (2005) information-structural distinction cannot satisfactorily predict the range of indefinite noun phrases that **can** occur in initial position. What we need then, is an account that provides a full and accurate explanation as to which types of indefinite noun phrases can occur as the initial NP of a specificational sentence and which cannot.

It is clear from examples such as (57) above, that indefinite noun phrases are not preferred as the initial NP of a specificational sentence. The flipside of this is that definite noun phrases are commonly found in *NP be NP* specificational sentences. I argue that this simple observation can provide us with the best starting point for finding out about the nature of specificational meaning and the structure of specificational sentences. Throughout the remainder of §3.3, I ask whether there is anything special about the semantics of definite noun phrases that makes them particularly suited to the specifying function and if, in fact, their unique semantics actually contributes to the creation of specificational meaning.

This leads me to argue for a semantic account of specificational sentences in which specificational meaning results from a nominal predication relation involving definite noun phrases; predicational (descriptive) and specificational (identifying)

meanings are understood as two possible interpretations of the same predication relation. In this analysis, specificational meaning is therefore not equated with syntactic movement and/or word order. Consequently, "reverse" specificational sentences can be accommodated into this analysis. Like Mikkelsen's (2005) inverse account, in my analysis the initial definite noun phrase of the "canonical" specificational sentence is non-referring. As a result, the account provided in this thesis can also accommodate the pronominalization data. However, rather than assuming that this NP moves from a predicative, postcopular position, I show that semantically, definite descriptions are inherently non-referring. Finally, by examining the differing semantics of definite and indefinite noun phrases, my analysis correctly predicts when an indefinite noun phrase will be permitted as the initial NP of a specificational sentence.

The analysis of specificational copular sentences provided in this thesis therefore has the advantage over alternative accounts in being able to accommodate and explain the full range of data outlined above. As my starting point, in the next section, I outline the Russellian analysis of definite descriptions upon which my semantic account of specificational *NP be NP* sentences depends.

3.3.2 The semantics of definite noun phrases

We can begin by looking at the differences between definite noun phrases and proper names outlined in Neale's (1990) Russellian analysis. The Russellian perspective assumes that definite noun phrases are 'object-independent' descriptions. As Neale (1990: 17) comments, "to know something by description it is not necessary to have had any sort of epistemic contact with the object that, in fact, satisfies the description one knows it under". Unlike 'object-dependent' expressions, such as proper names, definite descriptions are therefore not directly referring.

This analysis is supported by the behavior of definite descriptions in modal contexts. Let us assume for a moment that definite noun phrases are referring expressions and that (58a) is an identity statement. According to the Principle of Substitutivity, it follows that if two expressions refer to the same entity, then they can be substituted for one another without altering the truth value of the sentence. However,

since (58b) and (58c) below do not share the same truth value, this suggests that (58a) is not a genuine identity statement.

(58) (a) Yuri Gagarin = the first man in space (true)

(b) Yuri Gagarin might not have been Yuri Gagarin (false)

(c) Yuri Gagarin might not have been the first man in space (true)

(examples from Kearns 2000: 108)

We can explain these differing truth values as a consequence of the fact that modality deals with possible worlds. The reference of the proper name *Yuri Gagarin* is fixed; that is, it refers to the same individual in this or any other possible world. In contrast, the definite noun phrase *the first man in space* could describe different individuals in different possible worlds, in which another astronaut completed the mission before Yuri Gagarin. For this reason, the proposition in (58c) is true, leading us to conclude that definite noun phrases are descriptions rather than referring expressions.

Russell (1905) categorizes definite descriptions not with the object-dependent referring expressions, but as belonging to the category of quantified phrases. In particular, definite descriptions involve proportional quantification, which means that, like the quantifiers *all* and *most*, the definite article (*the*) quantifies over and in relation to a set. Additionally, definite descriptions have their own peculiar characteristics, such that they always expresses a proportion of a restricted rather than a general set; they involve universal quantification, meaning that the proportion of the set they express is the **whole** or **all** of it; and they have existential import (Neale 1990: 180).

To illustrate these properties, we can compare the behavior of quantifying expressions with *all*, which, like definite descriptions, involve universal quantification, but have neither of the other two characteristics. In (59), the proportion expressed is one of universal quantification over a very general set, 'all the children that exist'. However, sentence (60) must be understood as expressing universal quantification over a restricted set, such as 'all the children that live here'. Sentence (60) also assumes that the set

'children that live here' exists. This is not the case for sentence (59), which would be an acceptable proposition even in a world in which children ceased to exist.

(59) All children are naughty

(60) The children are naughty

Like other quantifying expressions, definite noun phrases are semantically descriptive rather than referring. This is not to deny that definite noun phrases can be used to pick out a particular individual, but, according to Grice (1969), such usage is a pragmatic, rather than a semantic, phenomenon. Grice makes a distinction between what a speaker **says** in the proposition they express and what a speaker **means** to communicate to their audience. For instance, by uttering sentence (61), a speaker may be expressing the Russellian proposition that 'all cats that live here are very timid'. However, they may intend to convey that 'a particular individual or referent is timid'. According to Grice, the speaker can communicate this secondary information by exploiting the assumption that both speaker and hearer are aware that a particular individual matches the description proposed in the definite noun phrase.

(61) The cat is very timid

As Neale (1990: 88) comments, "definite descriptions are particularly susceptible to referential usage because of their own particular semantics". Given that definite noun phrases involve an existential meaning, the members of the set expressed must exist. Furthermore, since the definite article expresses universal quantification over a restricted, rather than a general set, the number of individuals that exist which match this description is severely restricted. This is particularly true of definite descriptions involving singular nouns, where the set is so restricted that it has only one member. As Kearns (2000: 97) comments, "a singular definite description provides not only a familiar SET, but, more saliently, a familiar INDIVIDUAL" (emphasis original). However, from the Russellian-Gricean perspective, this usage is the result of shared

knowledge and is not part of the semantics of the proposition. While definite descriptions can therefore be "used" referentially, they are not directly referring. Instead, the referential usage of definite descriptions is derived via general pragmatic principles.

3.3.3 A semantic account of specificational meaning

In this section, I incorporate a Russellian analysis of definite noun phrases into a semantic account of specification. In §3.3.2, I showed that definite descriptions denote a universally quantified and existentially presupposed set. Here, I argue that this property of definite noun phrases makes them ideally suited to occurring in specificational sentences and is in fact what provides the sentence with its specificational meaning. I begin, in §3.3.3.1, by examining how nominal predication works before showing how specificational meaning results from the reinterpretation of a nominal predication relation involving definite noun phrases. At the outset then, "reverse" specificational sentences are accounted for in my analysis. In §3.3.3.2, I extend this account to "canonical" specificational sentences. I claim that this construction involves mismatch, whereby the semantic predicate is also the syntactic subject. This explains Mikkelsen's (2005) pronominalization data, outlined in §3.3.1. Finally, in §3.3.3.3, I show how my account of specificational meaning differs from and improves on alternative analyses. I explain that my account does not rely on the referential/attributive distinction, syntactic movement or a fixed word order, or hyponymy relationships and, as a result, does not suffer from the same problems as other, similar, accounts.

3.3.3.1 Predicate nominals and reverse specificational sentences

In construction grammar, predication involves a complex interaction of semantic and morphosyntactic information; if we take standard predication involving an active finite verb, the relevant semantic participant is linked to the referent of the verb's morphosyntactic subject. Predication, then, requires us to understand the semantics of different kinds of predicates, and not to view it solely as subject assignment. Verbs are understood to be inherently relational; that is, they presuppose the existence of another entity (see Langacker 1987: 304; Croft 1991: 62-63; Croft and Cruse 2004: 280-281).

For example, the verb *dances* requires a *dancer*. Proper names, on the other hand, are non-relational since they do not presuppose any other concept. Therefore, the sentence *Sarah dances* includes a relational predicate containing a 'substructure' which is elaborated by the non-relational argument *Sarah*.

However, nominal predication behaves differently from verbal predication. As Croft (1991: 63) notes, prototypical nouns, such as *man* are not relational, since "the existence of a man does not imply the existence of another entity". Furthermore, Francis (1999a, b) argues convincingly that even nouns which are inherently relational (including deverbal nouns) cannot assign thematic roles outside of their own noun phrases, *contra* Moro (1997). This begs the question then, what is the relationship between a predicative NP and its precopular subject?

Most authors agree that nominal predication involves a classifying relationship, whereby the subject referent is categorized as a member of a class or kind. For instance, (62) states that *Sally* is a member of the set *student*. This has the effect of ascribing a property to, or **describing**, *Sally*.

(62) Sally is a student [class-membership: predicational]

But how does this classifying predication relation come about? Indefinite noun phrases, such as *a student*, typically denote individuals; that is, they refer to a single **member** of a set (see §3.3.4). Therefore, it must be the **construction** which provides the postcopular NP with a predicative semantics via coercion (see the discussion on mismatch and predicate nominals in §2.4). In addition, it is the construction that imposes the syntactic relation such that the postcopular NP requires a syntactic subject. In (62) then, the predicate nominal construction imposes a class-membership predication relation onto the two noun phrases.[10] This allows the indefinite NP *a student* to be construed predicatively as denoting a set or kind.

[10] Francis (1999b) agrees; she claims that it is the construction which assigns the subject the "non-traditional" thematic role of 'classified' or 'that which is classified'.

I argue that sentences with definite NP predicates involve the same **classifying** predication relation as those with indefinite NP predicates, *contra* Croft (1991) and Francis (1999a, b).[11] From the Russellian analysis of definite descriptions, outlined in §3.3.2, definite noun phrases are understood to denote restricted sets. Consequently, example (63) can also be analysed as involving a classifying relationship; that is, the referent *Sally* is classified as a member (albeit the only member) of the set *best student we have*.

(63) Sally is the best student we have [class-membership: predicational]

Again, this classifying relationship allows for a "predicational" interpretation, in which an individual is ascribed a property. For example, in a context where the speaker is introducing *Sally*, example (63) **describes** her as *the best student we have*.

However, definite noun phrase predicates create a special type of class-membership predication relation. For example, in (63), the set *the best student we have* is so restricted that the referent *Sally* makes up the complete membership of this set. Consequently, this class-membership relation can be used not only to describe the referent, but also to identify or **specify** the list of members in the set, shown in (64).

(64) SALLY is the best student we have [class-membership: reverse specificational]

On this account, specificational meaning is therefore tied to the inherent semantics of definite descriptions, which denote **restricted** sets. The specificational interpretation of this class-membership relation is dependent upon a particular information-structure. For

[11] Croft (1991: 69) claims that sentences containing postcopular definite NPs involve "token-to-token identity" linked by an "equational *be*" or a "*be* of identity". Similarly, Francis (1999b) argues that such sentences are instances of a separate type of predicate nominal construction: the "identifying construction". She claims that this construction imposes an identifying relationship onto the subject and the postcopular NP, assigning the theta-role 'identified' to the subject. An important advantage to my account is that sentences with indefinite and definite noun phrase predicates are given a unified analysis as involving the same class-membership predication relation. From this, it follows that reverse specificational sentences do not require the stipulation of a separate type of predication relation or an equative semantics.

instance, in (64), the referent *Sally* is marked as focal; the result is a reverse specificational sentence.

Reverse specificational sentences therefore form a sub-construction of the larger predicate nominal construction (see Figure 3.2). I analyse the predicate nominal construction as involving a subject, a (semantically empty) copular verb and a predicative NP complement slot.[12] While reverse specificational sentences inherit this structure, they require a definite NP predicate. They also exhibit the same type of classifying predication relation as other instances of the predicate nominal construction, although the semantics of definite descriptions ensures that the set (or classifier) is restricted. However unlike the predicate nominal construction, reverse specificational sentences also have an additional information-structural requirement. In order for this predication relation to be interpreted specificationally, the referring expression must be marked as focal.

Evidence for analysing the reverse specificational construction as a subtype of the predicate nominal construction is provided by the diagnostics for predicative complement status. For example, whether the initial subject is in focus (creating a specificational interpretation) or not (creating a predicational interpretation), the subject-predicate string *Lucy the leader* is perfectly grammatical both in the absolutive *with XY* construction (see (65) and (66)) and embedded under *consider* (see (67) and (68)). This is not expected on an equative account, which analyses reverse specificational sentences as involving two referring expressions.

(65) A: Should we make Lucy the leader or Sarah?
B: Well, with LUCY the leader, the mission would surely fail

(66) A: Should we make Lucy the leader or the scout?
B: Well, with Lucy THE LEADER, I think we'd be in trouble

[12] I label the verb *be* as "semantically empty" because specificational meaning is produced by the relationship between the two NPs rather than being a lexical property of *be*. However, this is not to say that the copula is unimportant to the specificational construction. It is one of limited number of verbs that permit a specificational meaning relation to occur (see Higgins 1979: 161). However, since my focus is on *it*-clefts and related copular constructions, I do not discuss this issue in any more detail here.

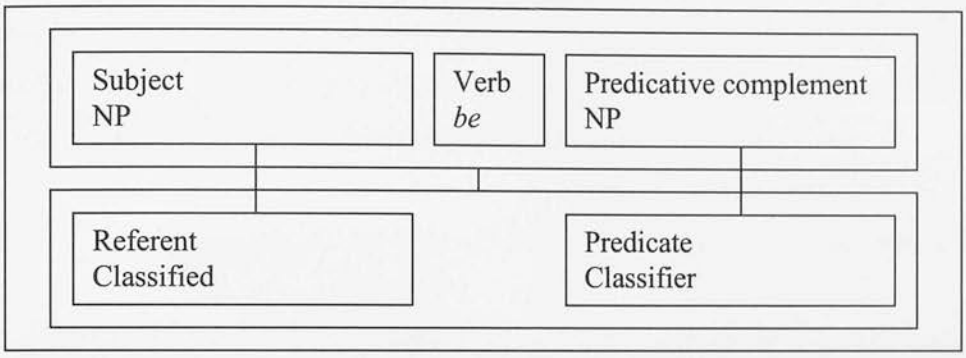
(67) A: Who is the leader?

B: I consider LUCY the leader

(68) A: What is Lucy's role?

B: I consider Lucy THE LEADER

The predicate nominal construction



The reverse specificational construction

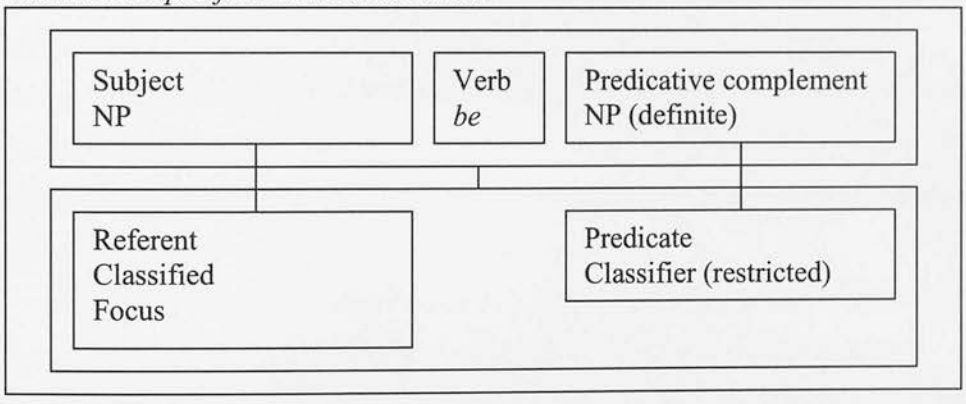


Figure 3.2 *Predicate nominal constructions*

Therefore, on my semantic account, specificational meaning involves a reinterpretation of the same class-membership predication relation which characterizes nominal predication more generally. Consequently, my analysis is able to accommodate reverse specificational sentences neatly: it explains how these sentences acquire a specificational interpretation while at the same time recognizing that they are essentially predicate nominal sentences. This represents an important improvement over the syntax-

based inverse analyses of specificational sentences. I come back to this issue in §3.3.3.3. In the next section, §3.3.3.2, I extend my analysis to canonical specificational sentences.

3.3.3.2 Canonical specificational sentences (and equatives)

In §3.3.1, I claimed that the subjects of "canonical" specificational sentences are semantically non-referring, based upon Mikkelsen's (2005) finding that such sentences pronominalize with *it* (shown in (69)). In what follows, I show how, once we apply a Russellian analysis of definite descriptions to this data, the non-referring semantics of specificational subjects can be given a simple explanation.

(69) The best student we have is Sally, isn't it? [class-membership: specificational]

Following a Russellian analysis, in (69) above, the definite NP subject *the best student we have* denotes a universally quantified and restricted set. Gricean reasoning has not been employed to convert this description into a referring expression (see §3.3.2). Mikkelsen's (2005) pronominalization data therefore falls out from my analysis, which views definite noun phrases as inherently non-referring. On this semantic account, there is no need to suggest that the initial NP moves from an underlying predicative position in order to obtain a non-referring semantics (see §3.3.3.3).

I argue that canonical specificational sentences, such as (69) above, involve the same class-membership relation as predicate nominal sentences with postcopular definite NPs (see §3.3.3.1). The only difference is that the restricted set (the predicate) is given prior to the identification of its members (the referent(s)). However, this word order only permits a specificational interpretation; it cannot be given a "predicational" (or descriptive) reading. For instance, in answer to the question posed in (70), the sentence in (70a), which contains an initial focused predicate, is unacceptable. Instead, the predicate nominal construct in (70b) is strongly preferred.

(70) Who is Sally? Tell me about her.

a) *The best student we HAVE is Sally.

b) Sally is the best student we HAVE.

In (69) above then, the class-membership predication relation is interpreted specificationally, as identifying the complete membership of the set *the best student we have*. Therefore, in sentences with this word order, the postcopular expression is always in focus (which may or may not be marked by intonation). This configuration conforms to (and is therefore motivated by) highly general information-structure tendencies. For example, there is a well-known preference for linguistic patterns in which given information is presented prior to new information and focal elements occur clause-finally (Ward, Birner and Huddleston 2002: 1372).

However, what is less clear is how the class-membership predication relation in canonical specificational sentences actually works. In §3.3.3.1, I claimed that indefinite noun phrases need to be in the predicate nominal construction (which imposes a classifying relationship on the subject and postverbal NP) in order to be construed as the semantic predicate. However, this is not required for definite noun phrases, which (according to a Russellian analysis) are semantically set-denoting. Indeed, there is an argument for suggesting that definite noun phrases are, in some sense, relational. As quantified phrases, definite descriptions have their own unique set of properties such that they denote complete (universally quantified) and restricted sets and have existential presuppositions (see §3.3.2). This means that definite noun phrases imply the existence of another entity – the membership of the existentially presupposed set.[13]

If we analyse the canonical specificational structure as consisting of a subject, a copular verb and a complement (which is not predicative), then a semantic predication can be derived compositionally from the semantics. From a constructional perspective, predication (or valence) is relative. As Croft and Cruse (2004: 281) comment, "predicate

[13] This is also what allows a definite description to be so easily converted into an act of reference (see §3.3.2). On this usage, the same NP that denotes a restricted set is used to refer to its members. In specificational sentences, on the other hand, the restricted set and its members are represented in two different NPs.

and argument status depend on what two semantic structures are being compared". In the case of the canonical specificational structure, a definite description is compared with a referring expression. This allows a predicational class-membership relation to ensue.[14]

On this analysis, canonical specificational and equative structures can be given a unified account. Both of these constructions have the same syntactic structure (subject-verb-complement). Unlike in the predicate nominal construction, this complement is not marked as predicative and so does not require a syntactic subject. This explains why canonical specificational and equative structures cannot occur in the *with XY* construction (see (71) and (72)) or embedded under *consider* (see (73) and (74)).

(71) A: Should we make Lucy the leader or Sarah?
B: *Well, with the leader LUCY, the mission would surely fail

(72) A: Is Cicero Tully?
B: *Well, with Cicero Tully…

(73) A: Who is the leader?
B: *I consider the leader LUCY

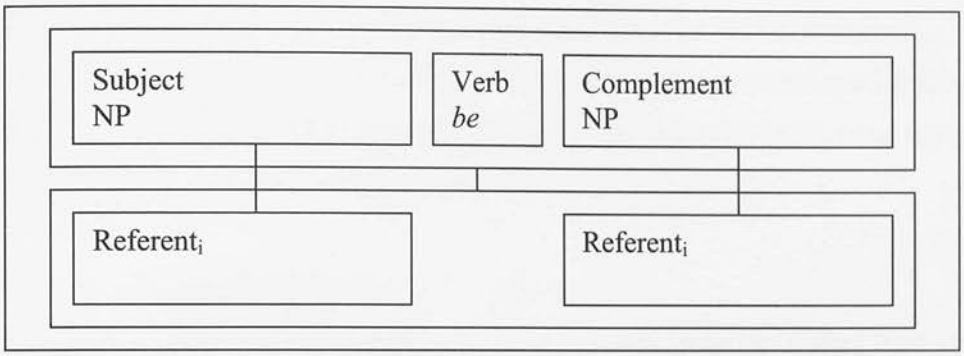
(74) A: Who is Cicero?
B: *I consider Cicero Tully

Since this structure is underspecified for predication, the semantics is not read off from the syntax. In equative sentences (where the subject and the complement are both semantically referring), a class-membership predication relation is not possible since referring expressions are not relational (see §3.3.3.1). However, in canonical specificational sentences, the subject denotes a restricted set, which, when paired with a referential complement, is understood to be the semantic predicate. In the canonical

[14] Therefore, while in predicate nominal sentences, the predication relation is constructional, in canonical specificational sentences it is compositional, deriving from the inherent semantics of definite noun phrases. From this, it follows that the predication relation in reverse specificational sentences is both constructional (inherited from the predicate nominal construction) and compositional (due to the semantics of definite NPs).

specificational construction then, we have a mismatch between the semantics and the syntax; that is, the initial NP is both the syntactic subject and the semantic predicate (see Figure 3.3).[15]

The equative construction



The canonical specificational construction

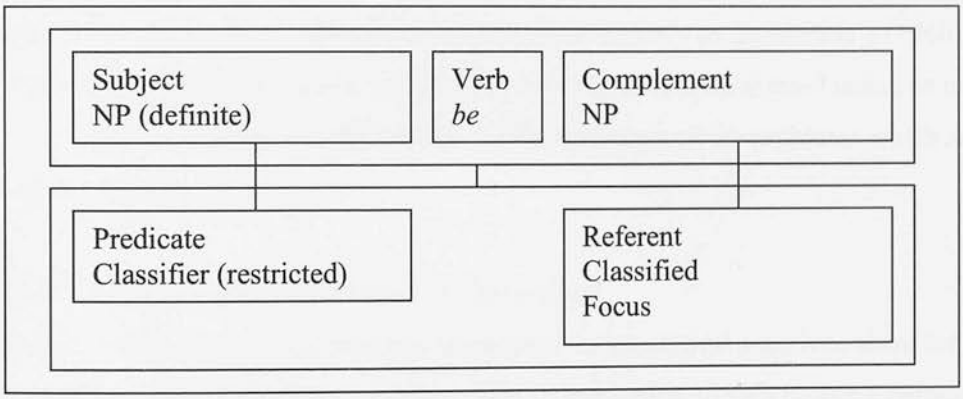


Figure 3.3 *Equative and canonical specificational constructions*

---

[15] Evidence for analysing the initial semantically predicating NP (rather than the postcopular referring expression) as the syntactic subject is provided by diagnostics for subject status. For instance, it is the initial NP that undergoes subject raising (see (i)) and is affected by the constraints on subject extraction when adjacent to complementizers such as *that* (shown in (ii)). This is also true of the initial NP of equative sentences.

(i) The murderer is John
a) The murderer is likely ____ to be John
b) *John is likely the murderer to be____

(ii) We all believe that the murderer is John
a) *It's the murderer that we all believe that ____ is John
b) It's John that we all believe that the murderer is ____

Although economy is not a priority for a constructional analysis, a motivated account, in which constructions are related within a hierarchical network, is strongly preferred. In the analysis developed in §3.3.3.1, reverse specificational sentences were shown to inherit the syntactic and semantic structure of the "classifying" predicate nominal construction. The construction's more idiosyncratic properties fall out from the inherent semantics of definite descriptions and additional information-structural requirements. Likewise, in this section, I have argued that the canonical specificational construction shares its syntactic structure with equatives, but exhibits the same semantic relation as the reverse specificational construction.

3.3.3.3 What my analysis isn't

In this section, I clarify how my analysis of specificational sentences differs from and improves on some of the alternative accounts proposed in the literature. I explain that because my analysis of specificational meaning does not rely on Donnellan's (1966) referential/attributive distinction, on syntactic movement or a fixed word order, or on a particular type of hyponymy relationship, it sidesteps many of the problems which affect otherwise similar accounts.

3.3.3.3.1 On the referential/attributive distinction

In §3.3.3.2, I argued that canonical specificational sentences and equatives share the same "underspecified" syntactic structure; that is, one which doesn't coerce a particular semantics. This analysis reflects the strong superficial similarity between these two sentence-types, as shown in (75) and (76).

(75) The guy at the bar is JOHN PRICE, isn't it? [specificational]

(76) The guy at the bar IS John Price, isn't he? [equative]

The difference in meaning between these two examples is entirely dependent upon whether Gricean reasoning, outlined in §3.3.2, has gone through. For example (76), we can imagine a context in which the speaker knows of two individuals (*the guy at the bar*

and *John Price*) but does not know that they are in fact the same person. In such a context, Gricean reasoning converts the initial definite description into an act of reference, as shown by the gender-specific pronoun *he*.

Nevertheless, the distinction between describing a set and referring to an individual cannot be reduced to the distinction between "unknown" and "known". As I explained in §3.2.3.2, Donnellan (1966) classifies definite noun phrases as either 'referential' (referring to a known individual) or 'attributive' (referring to an unknown individual). However, since even attributive definite noun phrases pronominalize with gendered pronouns, shown in (77), both types are treated here as an act of reference.

(77) The murderer, whoever he is, is insane, isn't he?

Regardless of whether the speaker is referring to an individual in the "real world" or whether the are referring to an assumed individual that exists only as a "mental" or "imagined" referent, they are taking the same step from describing a set to referring to an individual, known or unknown.

Therefore, the pronominalization data suggests that the referential/attributive distinction does not bring about a constructional difference (see also Donnellan 1966: 297). This represents an important problem for accounts of specificational meaning which rely on the referential/attributive distinction in order to distinguish between specificational and equative constructions, such as Declerck (1988) (see §3.2.3.2). Instead, I argue that the relevant constructional distinction between the canonical specificational and equative sentences in (75) and (76) above is that while the latter involves a 'referential reading' of the subject noun phrase, the former involves a 'description reading', in which the definite NP denotes a set, but is not used to refer to an individual. While definite NP subjects of equative (and also predicational) sentences may be referring to different degrees of precision (shown in (77)), in such cases Gricean reasoning has been employed, which effectively converts a description into an act of reference.

3.3.3.3.2 On syntactic movement and a fixed word order

The constructional account sketched over §3.3.3.1 and §3.3.3.2 shares a number of properties with Mikkelsen's (2005) inverse analysis of specificational sentences (outlined in §3.2.2). For one thing, it retains the important intuition that specificational sentences are not semantically equative and that the asymmetric relation between the two NPs is a case of apparent predication. Like Mikkelsen, I treat the initial NP of the canonical specificational structure as a non-referring, predicative element. Therefore, as I explained in §3.3.3.2, my semantic account is just as successful at accommodating and explaining the pronominalization data as Mikkelsen's (2005) inverse analysis. Finally, both Mikkelsen and I assume that the canonical specificational sentence type involves mismatch; that is, the initial NP is both the syntactic subject and the semantic predicate.

However, on the inverse account, this mismatch phenomenon is explained as a consequence of syntactic movement; the assumption being that mismatch configurations are derived from underlying structures which behave in expected ways. In contrast, in construction grammar, mapping between form and meaning is internal to the construction. Consequently, in this model, mismatch constructions are not derived and are simply presented as containing information which overrides inheritance from more general patterns (see §2.4). However, while mismatch configurations are tolerated in construction grammar, they must be explained; that is, there must be some motivation as to why this particular construction should appear in the language. In the account given in §3.3.3.2, I claimed that canonical specificational sentences inherit the same (underspecified) syntactic structure as the equative construction. Furthermore, I argued that the class-membership predication relation of the canonical specificational construction follows from the inherent semantics of definite descriptions, which (on a Russellian analysis) denote a complete and restricted set and presuppose the existence of its membership. Finally, I claimed that this mismatch construction is motivated by highly general information-structure principles.

An important advantage to my, largely semantic, analysis is that it is able to explain the relationship between predicational (descriptive) and specificational (identifying) meaning. In §3.3.3.1, I showed that specificational meaning involves a

reinterpretation of the class-membership relation involving definite NP predicates. In contrast, Mikkelsen's (2005) inverse analysis is a syntax-based approach which does not examine the origins of specificational meaning. Nevertheless, since specificational sentences are defined as "inverse" predicational sentences, we must conclude that, on this approach, specificational meaning is the result of syntactic movement.

As I explained in §3.2.2, there are two important problems with this syntax-based account. First, no explanation is provided as to why a syntactic movement operation should lead to this change in meaning from description to specification. Secondly, this account assumes that specificational sentences must always have an "inverse" word order. This means of course that the inverse analysis cannot accommodate reverse specificational sentences. Despite their specificational (or identifying) meaning, Mikkelsen (2005: 72) categorizes reverse specificational examples as "predicational". On my own semantic account, on the other hand, specificational meaning is not equated with syntactic movement. While my analysis retains the intuition that reverse specificational sentences share the same predication relation and syntactic structure as other instances of the predicate nominal construction, it also explains how specificational meaning comes about (as the result of the inherent semantics of definite descriptions and properties of information-structure). Consequently, on my account, the reverse specificational data is both accommodated and explained.

3.3.3.3.3 On hyponymy relationships

The analysis of specificational sentences developed over §3.3.3 has a number of points in common with the account of Blom and Daalder (1977). These authors suggest that predicational copular sentences involve a class-membership relation in which the subject referent is a hyponym of the superordinate category indicated by the predicate nominal. They claim that specificational sentences have the reverse alignment, with the subject denoting a more general concept than the postcopular NP. However, Declerck (1988: 92) criticizes this account. He notes that although predicational copular sentences typically express a class-membership relation, "there are subtypes of predicational

sentences where no idea of class-inclusion (or class-membership) appears to be present" citing examples with definite noun phrase predicates such as (78).

(78) John is the best musician in town (Declerck 1988: 92)

He also argues that specificational sentences do not involve a hyponymy relationship, since "There is no difference in generality between the variable NP and the value NP" (Declerck 1988: 93). Declerck suggests that this is the reason why specificational sentences are reversible, because they exhibit no difference in specificity. Furthermore, he notes that specificational sentences cannot be "reverse" predicational sentences because of the existence of reverse specificational sentences.

The analysis presented in this thesis provides a solution to the discrepancy between these two arguments. On the one hand, like Blom and Daalder (1977), I assume that canonical specificational sentences involve a class-membership predication relationship and are therefore, in some sense, "reverse" predicational sentences. However, I also agree with Declerck (1988) that specificational sentences do not involve a hyponymy relation and that specificational meaning is not tied to a particular word order. Where my account advances on this literature is by showing that the lack of a hyponymy relation does not prevent a class-membership relation. For sentences with definite NP predicates, such as (78) above, there is, as Declerck comments, no difference in generality. However, what Declerck does not recognize is that the predicate still denotes a set. This property of definite descriptions allows the copular sentence to acquire both predicational (x is a member of the set Y) and specificational (x makes up the complete membership of the set Y) interpretations. Unlike Blom and Daalder's (1977) analysis, this account is not dependent upon word order.

3.3.3.4 Interim summary

Throughout §3.3.3, I have shown that in both (canonical and reverse) forms of specificational copular sentence, one of the two NPs denotes a universally quantified, restricted and existentially presupposed set. This characteristic, which is inherent to the

semantics of definite noun phrases, is crucial for creating specificational meaning. As I go on to reveal in the following section, this analysis provides a comprehensive means of describing, and providing an explanation for, the limited distribution of indefinite NP subjects in the canonical specificational construction.

3.3.4 Explaining the distribution of indefinite noun phrases

In §3.3.3, I presented a largely semantic, constructional account of specificational sentences. I showed that this analysis can account for Mikkelsen's (2005) pronominalization data as well as accommodating reverse specificational sentences. In this section, I focus on the third and final set of data outlined in §3.3.1: the behaviour of indefinite noun phrases. So far, I have only examined specificational copular sentences containing definite noun phrases and have presented an analysis of specificational meaning which relies on the inherent semantics of definite descriptions. Here, I show that this account not only provides a reason as to why indefinite noun phrases rarely occur as the initial NP of a canonical specificational sentence but also explains why some indefinite noun phrase subjects are more acceptable than others.

From the Russellian analysis outlined in §3.3.2, definite descriptions are shown to involve proportional quantification, quantifying over and in relation to a set. Specifically, they express a universally quantified, restricted set. This means that the proportion of the set they express is the complete set. While definite descriptions involve proportional quantification, indefinite noun phrases involve cardinal quantification. Like noun phrases introduced by *some*, *many* and numbers, such as *four*, indefinite descriptions express a quantity rather than a proportion. For instance, the noun phrase *a man* highlights one of the members of the set *man*. However, to understand this description, we do not need to know what the set *man* constitutes. The total number of members in the set is irrelevant for the description and there are no restrictions in place for how general this set can be.

Because indefinite noun phrases involve cardinal, rather than proportional quantification, they are unlikely to occur as the subject of a canonical specificational sentence. When occurring in subject position, indefinite noun phrases denote individuals

rather than sets and so are unlikely to be interpreted as a semantic predicate, inducing a class-membership relation over a postcopular referent. Indefinite noun phrases are normally construed as sets or kinds only when occurring in the predicate nominal construction. However, while definite noun phrases denote very restricted sets, indefinite noun phrase predicates typically express general sets. For example in (79), *John* is identified as a member of the set *doctor*. Here, a specificational interpretation of this class-membership relation is not possible, since *John* does not come close to identifying the complete membership of the set *doctor*. The analysis presented in §3.3.3, therefore provides an explanation as to why canonical specificational sentences with indefinite noun phrase subjects, such as (80), are usually ungrammatical.

(79) John is a doctor

(80) *A doctor is John

However, occasionally, specificational copular sentences are found with indefinite noun phrase subjects. Examples (81) and (82) are provided by Mikkelsen (2005: 155) and Declerck (1988: 21), respectively. The grammaticality of these examples suggests that there is something special about these particular indefinite noun phrases that allows them to be accommodated (or coerced) into the subject position of the canonical specificational construction. From the analysis of specificational sentences outlined in §3.3.3, we would expect them to share some properties with definite noun phrases. A useful way into the discussion then, is to see how these examples compare to similar sentences containing definite noun phrase subjects and to ask what alternative function specificational sentences introduced by indefinite noun phrases provide.

(81) A philosopher who seems to share Kiparsky's intuitions on some factive predicates is Unger (1972), who argues… (Mikkelsen 2005: 155)

(82) An example of a superpower is the Soviet Union (Declerck 1988: 21)

When we examine the function of these two examples, it appears that they are in fact instances of two separate specificational constructions. I suggest that while the example shown in (82) serves to express a quantity (*one*) of a restricted and existentially presupposed set, the primary function of examples such as (81) is to bypass the existentiality condition of definite noun phrases. I discuss instances of each type in turn, beginning with those that behave similarly to example (81).

Mikkelsen (2005: 155) observes that example (81) occurs in a discourse about *Kiparsky*. However, the existence of a set of *philosophers who share Kiparsky's intuitions* is not asserted in the prior discourse and cannot be assumed to be shared knowledge. Such examples therefore differ from specificational sentences containing definite noun phrase subjects. To illustrate this point, consider the examples in (83) and (84).[16]

(83) *Presupposed: A psychologist works at Scarsdale Hospital*
a) The psychologist that works at the hospital is Dr. Amy Schrute
b) The psychologist is Dr. Amy Schrute

(84) *Presupposed: There are several psychologists working at Scarsdale Hospital*
a) #A psychologist who works at the hospital is Dr. Amy Schrute
b) *A psychologist is Dr. Amy Schrute

From (83), we can see that specificational sentences with definite noun phrase subjects presuppose the existence of the restricted set. However, the indefinite specificational

[16] In these examples, I have chosen to use *that* to introduce the relative clause in the definite NP in (83) and *who* in the indefinite NP in (84). I do this simply as a response to Bolinger's (1977: 12) claim that while the use of *that* is associated with given information, *who* indicates that the relative clause expresses new information. For example, Bolinger argues that in (i), the hearer is informed of the person's availability, while in (ii) their availability is not at issue. Although the information-structural distinction between *that* and *who* is only a tendency, and cannot alone be used as a diagnostic of given/new information, it is likely that specificational sentences introduced by the indefinite article will prefer use of *who* rather than *that*, as in (81).

(i) ...someone else who I knew was available (Bolinger 1977: 12)
(ii) ...someone else that I knew was available (Bolinger 1977: 12)

sentence shown in (84) does not require that the existence of the restricted set of *philosophers working at the hospital* is established in the previous discourse context or by shared knowledge. In fact, as shown in (84a), this specificational sentence becomes unacceptable when occurring in this context. This suggests that, in contrast to (82) above, the function of the indefinite article in (84) is not to describe a single member of an existentially presupposed set.

Instead, I argue that the primary function of examples such as (81) and (84) is to allow a specificational class-membership relation even when the restrictive set is not existentially presupposed. By introducing a specificational sentence with an indefinite article, the speaker is able to assert the existentiality of the set without it having to be shared knowledge. That is, the sentence in (84) effectively tells us both that *there is a psychologist who works at the hospital* and that *Dr. Amy Schrute matches this description*. Likewise, example (81) asserts that *there is a philosopher who shares Kiparsky's intuitions* and that *the referent Unger matches this description*. When accommodated into the subject position of a canonical specificational sentence, indefinite noun phrases are therefore understood to denote an existentially asserted, restricted set and will be interpreted as the semantic predicate.[17]

Aside from these presuppositional differences, the definite and indefinite examples in (83) and (84) above differ as to whether the restrictive relative clause *who/that works at the hospital* can be omitted. While specificational sentences with

---

[17] However, since the indefinite article does not provide universal quantification over this set, there is no exhaustivity presupposition such that the referent is necessarily the only individual that can match the description in the initial NP. This becomes even clearer when we examine specificational sentences with plural noun phrases in initial position. The definite NP subject in (i) involves proportional quantification, denoting a complete set. From this, it follows that the three individuals listed in the postcopular position make up a complete list of the members in this set. However, in (ii), the plural set is not universally quantified and so the three individuals listed in the postcopular position are understood as making up **either** the full set **or** a subset of members. For this reason, the verb *include* is required rather than *be*. Despite this, sentence (ii) still expresses a specificational meaning; as Higgins (1979: 161) observes, other verbs than the copular permit "a similar meaning relation" (see also footnote 12). This begs the question as to why indefinite NP subjects are ever permitted in the copular construction. I suggest that the reason is because the initial set is marked as singular and so can only be matched to a single referent in any one act of specification.

(i) The psychologists working at the hospital are Amy Schrute, Carl Scott and Andrea Smith
(ii) Psychologists working at the hospital include (*are) Amy Schrute, Carl Scott and Andrea Smith

definite NP subjects frequently occur with unmodified head nouns, specificational sentences with indefinite NP subjects do not, shown in (80) and (84b) above. In what follows, I explain why the indefinite NP subjects of specificational sentences are restricted in this way.

As I noted in §3.3.2, definite noun phrases always quantify over and in relation to a restricted set. However, part of this restrictive set information is often left unexpressed. For instance, the incomplete description provided in example (60), repeated here for convenience as (85), may express the same proposition as that in (86), where the background set is recovered in the form of a relative clause.[18]

(85) The children are naughty

(86) The children that live here are naughty

Kearns (2000: 80) suggests that the reason why this information is frequently omitted is because, "the speaker or writer assumes, or presupposes, that the audience can identify the background set, either from general shared knowledge, or because the information has been given earlier in the discourse".

For definite descriptions then, a restrictive context is expected and is assumed when left unexpressed. However, from the analysis presented above, the specificational examples with indefinite NP subjects, shown in (81) and (84), are not expected to behave in this way. Since the restricted set is not presupposed to exist and the indefinite NP serves to assert the existence this set, the information in the relative clause cannot be left unexpressed. If the initial NP is not understood to denote a restricted set, it cannot fulfill the requirements of the specificational copular construction, shown in §3.3.3.2.

Moving on now to the specificational example given in (82) above, and repeated here as (87), I suggest that this indefinite NP serves not to assert the existence of a

[18] In this semantic approach to quantifier domain restriction, sentences (85) and (86) are syntactically distinct, but they express the same semantic proposition. It is not assumed that incomplete descriptions contain unexpressed restrictive material at some underlying level of syntactic representation (e.g. a relative clause). Such an analysis is dubbed the 'syntactic ellipsis' approach by Stanley and Szabó (2000), and is strongly criticized therein.

restricted set, but to express a quantity (*one*) of a restricted and existentially presupposed set. Although example (87) is introduced by the indefinite article, this type of indefinite specificational sentence is usually introduced by the number *one*. For instance, while (84a) above was shown to be unacceptable in contexts where the restricted set is presupposed, the parallel sentence in (88a), introduced by *one*, is acceptable as long as it is understood as corresponding in meaning to (88b), selecting a single member out of the existentially presupposed and restricted set of *psychologists working at the hospital*.

(87) An example of a superpower is the Soviet Union

(88) *Presupposed: There are several psychologists working at Scarsdale Hospital*
a) One psychologist that works at the hospital is Dr. Amy Schrute
b) One of the psychologists working at the hospital is Dr. Amy Schrute
c) ??One psychologist is Dr. Amy Schrute
d) One such psychologist is Dr. Amy Schrute

Again, as shown in (88c), this type of indefinite specificational sentence also becomes less acceptable when part of the restrictive set information is left unexpressed. The reason for this lies in the differing semantics of indefinite and definite noun phrases. While definite noun phrases express proportional quantification over and in relation to a restrictive set, indefinite noun phrases involve cardinal quantification; they express a quantity of a set without placing any restrictions on how general this set can be. Consequently, when occurring with an unmodified head noun, as in (88c), the number *one* is understood to select a single member out of the very general set of *all psychologists that exist*. Only when the indefinite noun phrase contains either restrictive information (as in (88a)) or overtly refers back to an earlier mention of the restricted set (as in (88d)), will the indefinite noun phrase be permitted as occurring in the specificational construction.

However, a few exceptions to this pattern are presented in (89) and (90) below. In these examples, the head noun is not modified by a restrictive relative clause; that is,

part of the restrictive set information (*examples of superpowers* and *approaches to the problem at hand*) is left unexpressed. I suggest that the reason for this is that the restrictive information is presupposed lexically by the head noun. For example, *an example* must be *an example of something*. Likewise, *an approach* must be *an approach to something*. If this is the case, then these indefinite noun phrase subjects are permitted because the head noun requires a complement that can only be fulfilled by incorporating restrictive information from the discourse context.[19]

(89) *Presupposed: The are several superpowers in the world*
a) A good example is the Soviet Union
b) One example is the Soviet Union

(90) *Presupposed: There are several possible approaches to the problem at hand*
a) A good approach is to enlist the help of local businessmen
b) One approach is to enlist the help of local businessmen

From this analysis, we have seen that specificational sentences with indefinite NP subjects perform two possible roles: they are used to either assert the existence of a restricted set or to express a quantity of a restricted and existentially presupposed set. From the data, it seems that indefinite noun phrases are normally permitted into the subject position only if they quantify in relation to an overt and restricted set, which is often presented in the form of a relative clause. I explained this data by appealing to the inherent semantics of indefinite noun phrases. Unlike definite noun phrases, indefinite NPs do not provide restrictions as to the generality of the set they quantify. Consequently, when this restrictive set information is not overtly provided in an indefinite NP, it will not normally be assumed unless specified by the lexical head or by

[19] It is interesting that unmodified indefinite NP subjects become more acceptable if the head noun depicts a set that is already limited in number. For example, (i), involving the more restricted set *superpower*, seems to be more acceptable that the very general noun *woman* in (ii).

(i) ??One superpower is the Soviet Union
(ii) *One woman is Susan

anaphoric reference (see above). Since specificational sentences involve a class-membership relation between a referring expression and the description of a restricted set, only indefinite noun phrases that overtly provide restrictive set information can be coerced into the subject position of the canonical specificational construction.

This analysis of specificational sentences with indefinite noun phrase subjects has a number of advantages over alternative accounts. For example, in defense of the inverse analysis of specificational sentences, Mikkelsen (2005) attempts to explain the behaviour of indefinite noun phrase subjects as a consequence of information-structure. In what follows, I show that this account cannot accommodate the full range of data. The upshot is that by assuming an inverse analysis of specificational sentences, we cannot explain the distribution of indefinite noun phrases.

In §3.3.1, I explained that the behaviour of indefinite noun phrases poses a problem for the inverse analysis of specificational sentences. On this account, specificational sentences are derived via a movement operation which places the underlying predicate in the precopular position. Moro (1997) assumes that movement of the predicative noun phrase into the subject position can occur freely. Consequently, since indefinite noun phrases often occur as the predicate of a predicational copular sentence, they should, in theory, be able to occur in the subject position of a specificational sentence. Ungrammatical examples such as (92) are therefore unexpected on this account.

(91) John is a doctor

(92) *A doctor is John

In order to account for the behavior of indefinite noun phrases, Mikkelsen (2005) goes beyond Moro (1997), providing limits as to when "inversion" i.e. the raising of the predicative noun phrase, can occur. She notes that the raising of the referential noun phrase to the subject position is the unmarked or most likely occurrence; "Other things being equal, the most referential DP occupies the subject position" (Mikkelsen 2005: 163). However, Mikkelsen argues that if the predicative NP contains discourse-old

information it may be understood as the topic of the sentence. Consequently, "the preference for the topic to be in subject position may override this default alignment", resulting in a specificational sentence (Mikkelsen 2005: 163).[20]

Although a discourse-old predicate does not force a specificational argument structure, Mikkelsen argues that the raising of the predicative noun phrase will only be permitted if this information-structure criterion is fulfilled.[21] By introducing this caveat, Mikkelsen (2005: 153) explains the restrictions on indefinite specificational subjects as a consequence of the association of indefinite noun phrases with discourse-new information. As a result, indefinite noun phrase predicates rarely meet the criterion for a verified topic, and do not qualify for the role of subject.

Mikkelsen (2005) suggests that indefinite noun phrases can only move into subject position if they contain discourse-old information.[22] This analysis can therefore accommodate examples such as (81), repeated here as (93). Mikkelsen notes that in this example, the indefinite noun phrase subject involves discourse-old information, since *Kiparsky* has been mentioned earlier in the text.

(93) A philosopher who seems to share Kiparsky's intuitions on some factive predicates is Unger (1972), who argues… (Mikkelsen 2005: 155)

---

[20] A potential problem with Mikkelsen's (2005) analysis is that she assumes that 'topic' and 'discourse-old' are equivalent notions. However, topics may be discourse new. For example, in (i), *that guy that keeps calling* is new to the discourse even though the individual described is known to the hearer.

(i) That guy that keeps calling, he turned up at my work today.

[21] Mikkelsen (2005) argues that a discourse-old predicate cannot force a specificational alignment, in which the predicate precedes the subject, because such an analysis would predict that sentences such as (i) could not occur. While Mikkelsen (2005: 177) labels this example 'predicational' in response to its alignment, such sentences are referred to as 'reverse specificational' sentences elsewhere in this thesis.

(i) A: Who is **the winner**? B. John is **the winner**.

[22] A similar explanation for the fact that specificational subjects are typically definite noun phrases is provided by Declerck (1988). Like Mikkelsen (2005), he suggests that indefinite noun phrases are used to introduce new information and only occur as the subject of a specificational sentence if they contain "modifiers expressing old information" (Declerck 1988: 19).

However, as Mikkelsen (2005: 159) herself acknowledges, even when indefinite noun phrase predicates do express discourse-old information they may still not be permitted as the subject of a specificational sentence, shown in (94b). Consequently, Mikkelsen's information-structure account does not sufficiently counteract the serious claims made against the inverse analysis for its inability to explain the behavior of indefinite noun phrases.

(94) (a) A. I'd love to marry **a doctor**.
B. John's **a doctor**! [predicational]
(b) A. I'd love to marry **a doctor**.
B. ***A doctor** is John! [specificational]

In contrast, the analysis proposed in this thesis can explain why some indefinite specificational sentences, such as (93), are grammatical, while others, such as (94b) are not. Mikkelsen (2005) is of course correct that definite noun phrases are associated with discourse familiar information. This follows straightforwardly from an analysis which views definite noun phrases as involving the universal quantification of an existentially presupposed set. However, while Mikkelsen claims that the example in (93) also contains discourse-old information, the set *philosopher who shares Kiparsky's intuitions* is not presupposed. As I explained above, specificational sentences with indefinite NP subjects therefore have a unique function: to assert (rather than presuppose) the existence of a restricted set at the same time as involving it in a class-membership relation.

Rather than requiring discourse-old or presupposed information, it seems then that the fundamental criterion for a (canonical) specificational sentence is that the initial NP must express a restricted set. Since indefinite noun phrases can express a quantity of a very general set, this restrictive set information must be overtly provided. This criterion seems to be much more consistent than the information-structure distinction proposed by Mikkelsen. Regardless of whether the subject, *a scientist*, is discourse-new or discourse old, (95) is an unacceptable specificational sentence. While Mikkelsen's

account cannot predict this, the semantic analysis presented here states that indefinite noun phrases require restrictive set information in order to occur as the subject of a grammatical specificational sentence, as in (96).

(95) *A scientist is Dr John Matthews

(96) A scientist who is currently working on the new drug is Dr John Matthews

An alternative explanation for the fact that indefinite specificational subjects require restrictive or modifying information has been proposed by Heycock and Kroch (2002). They comment that indefinite specificational subjects "are good only to the extent that they have a "specific" reading" (Heycock and Kroch 2002: 112). Since additional material is often said to facilitate a specific reading for indefinite noun phrases (see Fodor and Sag 1982), one could argue that the grammaticality of relativized indefinite specificational subjects is due to a specificity condition. In this analysis, *a scientist*, in (95) above, is considered to be a non-specific indefinite noun phrase; that is, the speaker has no particular scientist in mind. However, in the specificational sentence in (96), *a scientist that is currently working on the new drug* is a specific indefinite NP, since it can be argued that the speaker (although perhaps not the hearer) has a particular individual in mind.

However, there is an important problem with this explanation: specificity is linked to referential use. From the following pair of examples, we can see that the specific indefinite noun phrase subject of the predicational copular sentence in (97) agrees with the gender-specific tag question. However, this is not possible for the specificational sentence in (98).[23]

(97) A doctor who finds cures for rare diseases lives there, doesn't he?

(98) *A doctor who finds cures for rare diseases is John Almond, isn't he?

[23] However, anaphoric *it* is also not possible for example (98), as shown in the example below. The reason for this is that *it* is inherently definite and cannot refer to indefinite noun phrases.

(i) *A doctor who finds cures for rare diseases is John Almond, isn't it?

This suggests that when occurring in the specificational construction, indefinite noun phrases are treated as semantic predicates, describing a restricted set *doctor who finds cures for rare diseases*, rather than referring to an individual. Furthermore, since unmodified noun phrases can have both generic and specific readings, this account does not satisfactorily explain why sentences such as (95) above, are ungrammatical.

In this section, I have shown that as well as accounting for Mikkelsen's (2005) pronominalization data and reverse specificational sentences, my constructional analysis of specificational meaning can also explain why some indefinite noun phrases can be accommodated into the subject position of a canonical specificational sentence and provides restrictions as to when they are allowed to perform this role. In addition, this data, which shows that restrictively modified noun phrases are preferred over unmodified indefinite noun phrases, provides support for my analysis of specificational sentences, in which a restricted set (existentially presupposed or asserted) enters into a class-membership predication relation with a referring expression.

3.4 Chapter summary

I began this chapter by claiming that *it*-clefts should be examined in relation to other specificational copular constructions (see §3.1). However, as I explained in §3.2, there is currently no agreed definition of what a specificational sentence is or how it works. As a result, in §3.3, I built up my own analysis of specificational copular constructions. I began by examining the role of definite noun phrases in these sentences and adopted a Russellian analysis of definite descriptions. I claimed that specificational meaning results from a class-membership predication relation involving definite NP predicates. I showed that my analysis of specificational copular sentences is able to both accommodate and explain more data than the alternative accounts. In chapter 4, I apply this analysis of *NP be NP* specificational sentences to clefts. I suggest that from this perspective, our understanding of the behaviour and structure of *it*-clefts can be much improved.

4. A CONSTRUCTIONAL ACCOUNT OF CLEFT SENTENCES AS SPECIFICATIONAL STRUCTURES

In this chapter, I apply the constructional account of specificational *NP be NP* sentences from chapter 3 to *it*-clefts. This leads me to argue for a (non-derivational) extraposition account of *it*-clefts, in which the cleft clause is analysed as a restrictive relative, modifying the element *it*. I claim that this 'discontinuous constituent' functions like a definite noun phrase, allowing *it*-clefts to be analysed as involving a nominal predication relation, just like the *NP be NP* sentences in §3.3. I show how once we take this step, many of the *it*-cleft's seemingly idiosyncratic properties are inherited from the wider specificational construction and can be explained either as the result of the semantics of definite noun phrases or as a product of specificational meaning.

The chapter is structured as follows. In §4.1, I outline my analysis of *it*-clefts and go on to show what it buys us. In this section, I frame my discussion in response to the expletive analyses discussed in §3.1. Focusing in particular on the *it*-cleft's pragmatic properties and on the status of the cleft clause, I show that my analysis, which views clefts primarily as specificational copular sentences, both fits and explains the data far better. In §4.2, I explain how the particular analysis of *it*-clefts argued for in this thesis is superior to other extraposition accounts that have been proposed in the cleft literature. Here, I outline some fundamental problems with the earlier transformational accounts and show how my analysis advances on the recent 'discontinuous constituent' accounts of Percus (1997) Hedberg (1990, 2000) and Han and Hedberg (2008).

What is unique about the analysis argued for here is that it assumes that *it*-clefts and specificational copular sentences in general involve a nominal predication relation that is tied up with the semantics of definite noun phrases. This provides support for a discontinuous constituent analysis by explaining why *it*-clefts should contain a definite-like description in subject position. Furthermore, although other authors have identified a number of properties that *it*-clefts share with definite noun phrases, I explain how these properties come about by applying classic philosophical work on definite descriptions to the *it*-cleft data. Finally, in §4.3, I respond to some of the arguments that have been

raised against extraposition analyses of *it*-clefts. Here, I claim that these objections do not represent an insurmountable problem for the analysis proposed in this thesis and I show that much of the "difficult" data can be accommodated.

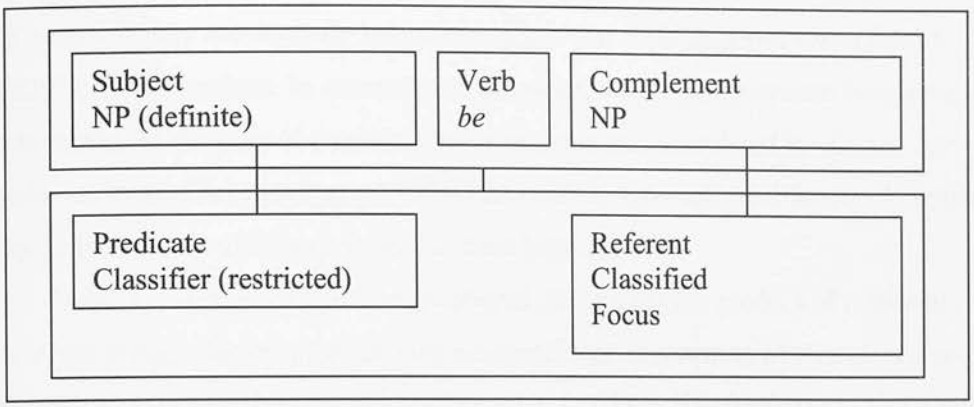
4.1 A motivated analysis of *it*-clefts

In this section, I describe an analysis of *it*-clefts which treats them as a subtype of specificational copular sentence. I argue that the initial *it* and the cleft clause together form a discontinuous definite description, with *it* functioning both like the definite article and as a semantically general head noun. The cleft clause is classified as an extraposed restrictive relative clause which modifies *it*. With this analysis in place, my account of specificational meaning from chapter 3 can be extended to *it*-clefts.[1] My claim is that *it*-clefts inherit from the wider (canonical) specificational copular construction. This is shown in the diagrammatic representation given in Figure 4.1. Like other specificational copular sentences, the *it*-cleft is a nominal predication structure containing a (discontinuous) definite NP predicate. The focused postcopular NP (the clefted constituent) refers to the complete membership of the set described by the definite noun phrase. In what follows, I go through the reasoning that brings me to this analysis of *it*-clefts and show that many of the *it*-cleft's functional and structural properties fall out from my analysis.[2]

[1] My discussion centres on *it*-clefts with NP foci (as shown in Figure 4.1). As I explained in §1.1, *it*-clefts also allow a variety of non-nominal elements into the postcopular position. However, in §5.3, I show that the range of possible *it*-cleft foci is not inherited from other constructions. Since the purpose of this chapter is to highlight patterns of correspondence between *it*-clefts and other specificational constructions, I do not discuss *it*-clefts with non-NP foci in detail here. These instances are given a diachronic explanation and are properly integrated into my analysis in chapters 7 and 8.

[2] As shown in Figure 4.1, my analysis effectively reduces idiosyncrasy in the *it*-cleft construction to two main properties: the restrictive modification of the pronoun *it* and the extraposition of the relative clause. These construction-specific properties are of course problematic for this analysis if they are not motivated (see §4.3) and I go on to provide a diachronic explanation for them in chapters 7 and 8.

The canonical specificational construction



The it-*cleft construction (with NP foci)*

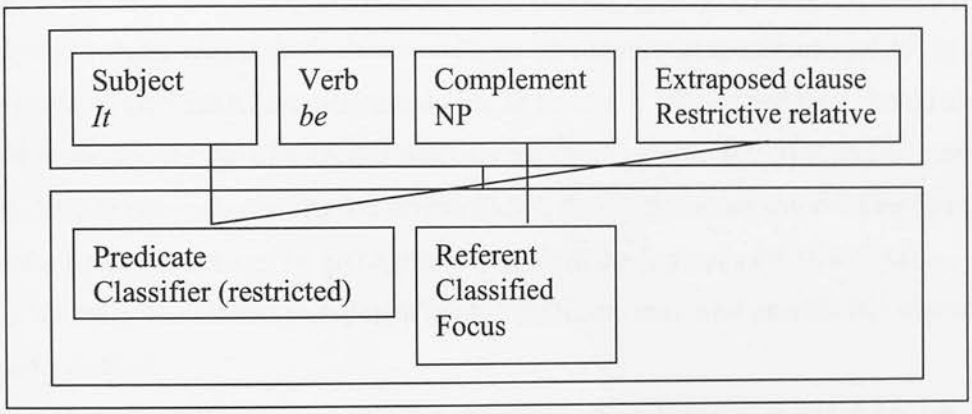


Figure 4.1 *The* it-*cleft construction and the overarching specificational schema*

4.1.1 A family of specificational copular constructions

In this section, I explain the reasoning behind my analysis of *it*-clefts. I start from the premise that *it*-clefts have a specificational meaning. For example, the *it*-cleft in (1) identifies *orange soda* as *my favourite drink*. As I noted in §3.2, *it*-clefts share this property with noncleft specificational *NP be NP* sentences and with the so-called 'pseudoclefts' (including *th*-clefts and *wh*-clefts).

(1) It's orange soda that I like best [*it*-cleft]

(2) What I like best is orange soda [*wh*-cleft]

(3) The thing that I like best is orange soda [*th*-cleft]

(4) My favourite drink is orange soda [noncleft specificational]

Since all of these examples share the same specifying function and (as copular constructions) they are formally related, I suggest that these sentence types form a "family" of constructions. In construction grammar, the language system is represented as a hierarchical network of constructions. Constructions are related to one another via inheritance links. I suggest that what links all of these types of specificational copular sentence is that they all inherit from the same basic schema.

In §3.3, I suggested that specificational meaning is the product of a special type of nominal predication relation. In such sentences, the description of a restricted set enters into a class-membership predication relation with a referring expression. Since definite noun phrases inherently describe a universally quantified restricted set, this analysis explains why definite descriptions are so common in specificational *NP be NP* sentences. If cleft sentences are indeed specificational, it follows that they should also involve the description of a set and that they may well exhibit properties in common with definite noun phrases. As we saw in §3.3.4, the restrictive set information found in definite noun phrases can be given, or recovered, in the form of a relative clause. Consequently, the clausal constituent in cleft sentences may well provide this restrictive set information.

With *th*-clefts, the comparison with specificational sentences introduced by definite noun phrases is easily made. *Th*-clefts are defined by Collins (1991b: 483)) as pseudoclefts "introduced by *the* in conjunction with the proform equivalents of the English interrogatives (*thing*, *one*, *place*, *time*, *reason*, *way*)". In other words, *th*-clefts, such as (5), are introduced by a precopular definite description containing a restrictive relative clause.

(5) [The [one [that left]]] was Howard [*th*-cleft]

Following a Russellian analysis of definite descriptions, we can therefore assume that the definite article quantifies proportionally over the restrictive set *one that left*, with the lexical item *one* performing the role of head noun, which the relative clause modifies.

As with canonical (i.e. non-reversed) *NP be NP* specificational sentences, the unique member of this restricted set (the referent *Howard*) is located in the postcopular position.

This analysis can be extended to *wh*-clefts, with the only difference being that the precopular constituent is a 'fused relative' construction. Ward, Birner and Huddleston (2002: 1420) show that fused relatives, such as that found in (6) below, correspond to definite noun phrases with integrated relative clauses: in this case '*the x [I like x best]*'.

(6) [What [I like best]] is orange soda

From this, it follows that the introductory element *what* performs the two roles of providing definite-like quantification as well as acting as the head noun of the description. As a result, the analysis of specificational meaning as involving a nominal predication relation can be applied to *wh*-clefts.

It-clefts do not show the same superficial correspondence to noncleft specificational sentences as *th*-clefts and *wh*-clefts. The reason for this is that the cleft clause is located in an extraposed sentence-final position. However, once we assume that the cleft clause is a restrictive relative (an assumption backed by considerable evidence presented in §4.3), the formal resemblance between *it*-clefts and specificational *NP be NP* sentences becomes apparent. The function of a restrictive relative clause is to modify a nominal antecedent. However, it is generally agreed throughout the cleft literature that the cleft clause does not restrictively modify the postcopular constituent.[3] So what then does it modify? The answer is that the cleft clause restrictively modifies, and forms a semantic unit, with the initial *it*.

(7) [It t_i] was Howard [that left]$_i$

[3] For example, when talking about his earlier extraposition (or "transposition") theory, Jespersen (1937: 83) notes that "the clause is felt to be, and is treated like, a relative clause, though it does not logically restrict the word with which it is connected".

Although *it* normally functions as a complete pronominal NP, I suggest that in *it*-clefts, this element performs a role equivalent to *the one* or *the thing* in *th*-clefts. The initial *it* and the cleft clause operate together like a discontinuous definite description; that is, the constituent *it* provides definite-like quantification as well as acting as the head noun which is restrictively modified by the cleft clause.[4]

I conclude then that all types of cleft sentence, including *it*-clefts, can be analysed as subtypes of specificational copular sentence.[5] In these examples, specificational meaning comes about as a result of a nominal predication relation involving definite NP predicates. While the definite description denotes a restricted set, the postcopular element provides the membership of this set, thereby accounting for the identifying meaning associated with these constructions. However, by effectively reducing cleft constructions to specificational *NP be NP* sentences, we are left to ask, what is the theoretical significance of the concept "cleft"? Should cleft sentences be grouped together as forming a construction distinct from so-called "noncleft" specificational sentences?

Where cleft sentences differ from other specificational sentences is that the definite description is headed by a semantically general or underspecified noun. For example, while noncleft specificational sentences, such as (8), involve meaningful

---

[4] Support for a unified analysis of cleft sentences comes from Lambrecht (2001: 469), who notes that the constituent *what* in *wh*-clefts is a "composite element combining in a single word the functions of the morphemes *it* and *that* which in IT clefts appear in discontinuous form". He draws evidence for this suggestion from French "where *what* is expressed by the sequence *ce que* or *ce qui* 'it that'". On this basis, Lambrecht argues in favour of an expletive analysis of cleft sentences in which the introductory items *what*, *it…that* and *the one that* are all semantically inert dummy elements (see also §6.2.1). However, the suggestion that *it* and *that* form a unit, corresponding to the composite element *what* provides equal support for the analysis presented here, in which the cleft pronoun and the cleft clause function together as a discontinuous definite description.

[5] In addition to *th*-clefts, *wh*-clefts and *it*-clefts, other types of cleft sentence have been noted in the literature. While these are also subtypes of specificational copular construction, they express meanings that go beyond or differ from definite-like quantification. As a result, I do not discuss these instances here. For example, *all*-clefts, which are classified by Collins (1991b: 483) as a distinct type of pseudocleft, have a "below expectation" scalar reading (see Traugott 2008). For example, the meaning of *all* in (i) corresponds to *the **only** thing*.

(i) All he cares about is himself

Furthermore, as I explained in §3.1, the *it*-cleft's structural configuration is shared by demonstrative clefts and *there*-clefts. While demonstrative clefts have an additional deictic meaning, *there*-clefts do not involve definite-like quantification, as shown in §4.1.2.3.

nouns, the corresponding cleft sentences are introduced by semantically general nouns such as *one* or *thing*, as shown in the *th*-cleft given in (9).

(8) The **bassoon player** is Howard [specificational noncleft]

(9) The **one** that plays the bassoon is Howard [*th*-cleft]

These nouns require modification in order to give a sufficiently informative description; without it they cannot identify a restricted set. For this reason, in cleft sentences, the relative clause, which provides restrictive set information, is obligatory, as shown in (10) and (11).[6]

(10) *The one was Howard

(11) *What is champagne

So-called "truncated clefts", such as (12), provide an apparent exception to the claim that cleft sentences always require restrictive modification. However, I suggest that these examples do not involve the "truncation" of an *it*-cleft, but instead are instances of a separate specificational construction, involving *it* as a full pronoun, which is anaphoric to a complete set description. Similar analyses are proposed by Hedberg (2000) and Mikkelsen (2007).

(12) A. Is John **the murderer**? B. No, **it**'s Gary.

Therefore, once we incorporate the different types of cleft sentence into a unified analysis of specificational copular constructions, the only property that distinguishes "clefts" from "nonclefts" is that, in the former, the precopular description is introduced by an underspecified or semantically general head noun which requires restrictive

[6] Lambrecht (2001: 469) uses the ungrammaticality of sentences such as (10) and (11) as evidence that *the one* and *what* are "dummy elements" rather than argument expressions. However, the analysis provided in this thesis explains why these examples are ungrammatical, without requiring that the introductory items are semantically redundant.

modification. It is not at all clear to me whether this feature is sufficient to collect or group these different cleft constructions together as forming an overarching cleft schema. For many authors, the concept "cleft" is inextricably tied to the correspondence of these sentences to their noncopular counterparts[7]. However, since I view these examples primarily as specificational copular sentences, the term "cleft" has no theoretical significance here[8]. As a result, it may be the case that the concept of "cleft sentence" has no psychological reality either. What is more, my unified analysis of specificational copular sentences raises the question of whether the *th*-cleft should be recognized as a distinct construction. Unlike *it*-clefts and *wh*-clefts, *th*-clefts are not formally distinct from noncleft *NP be NP* sentences. Since the semantic content of the head noun is a matter of degree, rather than an absolute distinction, it is not clear where the dividing line between them is and whether they truly represent two separate constructions.

In this analysis then, cleft sentences are treated as fully fledged copular constructions rather than as focusing devices which "cleave" the informational content of simple noncopular sentences. The *it*-cleft analysis formulated here therefore differs considerably from the expletive analyses outlined in §3.1 (including Jespersen 1937, Chomsky 1977, Williams 1980, Delahunty 1982, Rochemont 1986, Heggie 1988 and É. Kiss 1998). The expletive approach maximizes correspondence between *it*-clefts and their noncopular counterparts. As Hedberg (1990: 35) comments, this approach results in a counterintuitive analysis in which "nothing is what it seems": the constituent *it* is analysed as a meaningless expletive element, the postcopular constituent is really a preposed argument and although the cleft clause looks and behaves like a restrictive relative, it is treated as a non-modifying predicate.

---

[7] For example, Collins (1991b: 484) limits his examples of *th*-clefts to only those sentences that can be analysed "as in any sense a cleaving of a simple sentence". Likewise, Lambrecht (2001: 503) argues that the nonequivalence between proverbial *it*-clefts and noncopular subject-predicate sentences "strongly suggests noncleft status of this construction" (see §5.2).

[8] Ward, Birner and Huddleston (2002: 1423) note that the terms 'cleft' and 'pseudocleft' "serve as convenient metaphors" for the fact that while *it*-clefts almost always have noncopular paraphrases, *wh*-clefts and *th*-clefts do not always have simple sentence counterparts.

In contrast, I have argued that the initial *it* is not an inert element and in fact plays an important semantic function, providing definite-like quantification as well as acting as the head noun of the description of a set. As I go on to show in §4.1.2, once we recognize the quantifying function of initial *it*, we can provide an explanation for many of the *it*-cleft's pragmatic properties; without acknowledging the semantic contribution of this word, we cannot successfully account for the exhaustiveness and existentiality presuppositions in *it*-clefts.

In my account, the cleft clause is analysed a restrictive relative which modifies the initial constituent *it* but is in an extraposed sentence-final position. I assume that this property of the *it*-cleft is motivated by information structure principles, such as the tendency for heavy constituents to occur near the end of the clause (see Ward, Birner and Huddleston 2002: 1372). Although this analysis of the cleft clause is not uncontroversial, I provide further evidence for it in §4.1.3. I show that with this analysis in place, we can account for many of the *it*-cleft's (seemingly contradictory) properties without stipulating a special type of clause that is unique to the *it*-cleft construction.

While other authors in the cleft literature have advocated similar analyses to the one proposed here (see §4.2), what is truly unique about this analysis is that it is integrated into an original account of specificational sentences. I have argued that specificational meaning results from a nominal predication relation involving definite noun phrase predicates. This not only provides support for a discontinuous constituent analysis of *it*-clefts but it also explains why the definite-like description in *it*-clefts is so important to its specifying function. The upshot is a truly unified analysis of specificational copular sentences as involving the same type of nominal predication relation. This compares favourably to Heggie's (1988) attempts at integrating an expletive analysis of *it*-clefts into a unified account of copular constructions. As I explained in §3.1, Heggie assumes that the cleft clause in *it*-clefts is directly predicated of the clefted constituent and concludes that all copular constructions involve a predication relation of some sort. However, Heggie's expletive analysis does not provide the *it*-cleft with the same type of predication relation that exists in other specificational

sentences. As a result, she fails to recognize the role of nominal predication in creating specificational meaning for all such constructions.

4.1.2 Explaining the it*-cleft's pragmatic properties*

In this section, I provide support for the 'discontinuous constituent' analysis of *it*-clefts presented in §4.1.1. I show that many of the *it*-cleft's pragmatic properties, including focus, existentiality, exhaustiveness, contrast and givenness, are inherited from the wider specificational construction and from definite noun phrases in general. While others have recognized such similarities between the behaviour of definite noun phrases and *it*-clefts (see §4.2), here I provide an explanation for them in relation to my own account of specificational sentences, which involves a Russellian analysis of definite descriptions.

I conclude that the extraposition-from-NP analysis argued for in this thesis therefore has more explanatory power than the expletive analyses outlined in §3.1. In these syntax-centred approaches, the pragmatic properties of *it*-clefts are often regarded as peculiar to clefts. For instance, Delahunty (1984: 73) suggests that they lie "in the nature of cleft focusing itself". An important advantage to the analysis proposed here is that it is able to explain the *it*-cleft's semantic and pragmatic properties by appealing to more general linguistic patterns.

4.1.2.1 Focus

The *it*-cleft is widely regarded as a focusing construction; that is, the primary informational content is placed in the syntactically marked postcopular position and is typically coupled with prosodic prominence, as shown in (13).

(13) It was HOWARD that left

For proponents of the expletive analysis, the *it*-cleft's main function is to mark focus syntactically, providing an information-structure variant of corresponding simple noncopular sentences. However, in the analysis proposed here, focus in *it*-clefts is shown to be due to inheritance from the wider specificational copular construction.

I explained in §3.3.3 that specificational meaning in copular sentences results from a special type of predication relationship involving definite noun phrases. This same predication relationship allows for two possible interpretations: a predicational reading in which the referent is described as having a particular property and a specificational reading in which the referent provides the complete membership of a restricted set. The specificational reading is selected if either the referent is marked as focal by intonation, as in (14), or the referent is placed after the verb in a position that, in English, tends to contain the focal element, as shown in (15).

(14) JOHN is the murderer

(15) The murderer is John

From the premise that *it*-clefts are specificational copular sentences, it follows that they should inherit this property (see Figure 4.1 above). As a result, I claim that *it*-clefts do not function primarily as a focusing device, operating as an alternative to noncopular sentences. Instead, their main function is to provide specificational meaning. Since specificational meaning involves the listing of the members of an established set, focus will inevitably fall on the referent or referents that make up the membership of this set.

4.1.2.2 Existential presuppositions

A further property of the *it*-cleft is that it is presuppositional; that is, the information given in the cleft clause is not part of the assertion. For example, in sentence (16), the proposition that *someone left* is taken for granted as a precondition to the assertion that *Howard* was the one who left. Presuppositions are preserved under negation. For instance, in (17), we are told that *Howard didn't leave* but the presupposition *that somebody (else) did leave* remains in force (see Halvorsen 1978). Tellingly, this same property is also found in sentences containing definite descriptions. For example, the sentences in (18) and (19) both presuppose that *the murderer* exists.

(16) It was Howard that left

(17) It wasn't Howard that left

(18) The murderer is John

(19) The murderer isn't John

I argue that for all types of specificational copular sentence, existential presupposition can be explained as a consequence of the definite-like elements contained within them. As I showed in §3.3.2, definite descriptions involve existential commitment. That is, by using a definite noun phrase the speaker assumes that the described set exists. Consequently, any analysis of the presuppositional nature of cleft sentences must apply equally to definite descriptions.[9]

4.1.2.3 Exhaustiveness

In addition, it is often acknowledged that *it*-clefts are exhaustive. For instance, the noncopular subject-predicate sentence given in (20) does not preclude the possibility that additional people besides *Howard* also left on that occasion. However, from the corresponding *it*-cleft presented as (21), we assume that *Howard* is the only person that left. This property of exhaustiveness is also found in specificational *NP be NP* sentences containing definite descriptions. For example, (22) suggests that there is only one bassoon player in the band and that this description uniquely applies to *Howard.*

(20)	Howard left	[noncopular sentence]
(21)	It was Howard that left	[*it*-cleft]
(22)	The bassoon player was Howard	[specificational *NP be NP*]

[9] The concept of presupposition plays an important role in arguments about where the boundary lies between semantics and pragmatics (Huang 2007). As a result, exactly how the existential presupposition of definite noun phrases should be analysed is a matter of current debate. Some authors attempt to reduce the concept of presupposition to 'semantic entailment' for positive sentences and 'conversational implicature' for negative sentences (see Atlas and Levinson 1981 for this type of analysis of *it*-clefts). It is beyond the scope of this thesis to review these arguments here. Consequently, I continue to use the cover term 'presupposition' for meanings that remain even when the sentence is negated.

I suggest that the exhaustiveness in *it*-clefts and other specificational copular sentences originates from the semantics of definite descriptions. In §3.3.2, I noted that definite noun phrases involve universal quantification. This means that they express a complete set, encompassing all of its members.[10] Specificational sentences have an identifying function: when we identify the referents that a universally quantified expression describes, we necessarily get a complete list of all of the members of that set. This analysis is supported by Higgins (1979: 150), who claims that in specificational sentences, the postcopular noun phrase "says what constitutes or makes up the object [which for us is the set] referred to by the subject noun phrase". As a result, once we treat the initial *it* and the cleft clause of *it*-clefts as a definite-like description, the property of exhaustiveness falls out neatly.

This analysis contrasts with the assumptions of Atlas and Levinson (1981: 29), who argue that "there is no exhaustiveness presupposition for clefts". These authors suggest that positive clefts, such as (23), **entail** but do not presuppose that *Mary kissed exactly one person*. They claim that this proposition follows as a consequence of the fact that the example in (23) lists only one person (*John*). This fact, accompanied by the presupposition that *Mary kissed someone* entails that *Mary kissed exactly one person*. However, for negative clefts, Atlas and Levinson claim that exhaustivity is neither entailed nor presupposed. For example, they argue that because the singular *John* can be replaced by the plural *Mart and Rick* in (24), there is no conventional implication that *Mary kissed only one person*. They note, "What is being contradicted here is not a "presupposition" but an assertion, and there is no problem of felicitousness" (Atlas and Levinson 1981: 54). For Atlas and Levinson then, the positive cleft in (23) entails that *Mary kissed only one person*, but this is not retained for the negative sentence in (24). Consequently, they argue that there is no exhaustiveness implicature in either sentence.

(23) It was John that Mary kissed (Atlas and Levinson 1981: 21)

[10] As Hawkins (1978: 159) comments, definite noun phrases have "inclusive reference"; that is, they "refer to the totality of the objects or mass in the relevant shared set".

(24) It wasn't John that Mary kissed – it was Mart and Rick

(Atlas and Levinson 1981: 25)

An alternative analysis is provided by Declerck (1988: 30). He argues that the positive cleft in (23) has an exhaustiveness implicature stemming from the Maxim of Quantity ('Make your contribution as informative as required'). If the speaker adheres to this principle, they will provide an exhaustive list of the values that satisfy the variable in *Mary kissed x*. He notes, "It is clear, then, that exhaustiveness follows directly from the act of specification itself. Exhaustiveness is nothing else than 'exhaustive listing'" (Declerck 1988: 30). Like Atlas and Levinson (1981), Declerck (1988: 31) assumes that negative clefts do not have an exhaustiveness implicature. He argues that examples such as (24) are not really specificational, since a value is not specified for a variable; without the act of specification, exhaustiveness is not implied. For Declerck (1988) as well as Atlas and Levinson (1981) then, exhaustiveness may be associated with positive clefts (either by entailment or implicature) but is not present in negative clefts.

However, in my analysis, exhaustiveness is a consequence of the universal quantification involved in definite descriptions; it is not necessary to stipulate that it is implicated or entailed by the act of specification itself. Therefore, the implication of (23) is not that *Mary kissed only one person* but that *John* makes up the complete membership of the set of individuals that Mary kissed. Consequently, (24) is not problematic, since all that is asserted is that the entire set of people that Mary kissed is not made up of *John* but of *Mart and Rick*. Atlas and Levinson are therefore correct that (24) involves the contradiction of an assertion rather than a presupposition. However, the universal quantification of the definite-like description is not part of this assertion and is consequently retained in the negative statement. Therefore, 'exhaustiveness', or the fact that we are talking about the complete set of people that Mary kissed, is an equally prominent feature of both positive and negative clefts.

The situation is only slightly different for *NP be NP* specificational sentences. Unlike in *it*-clefts, the precopular definite NP of these sentences contains a head noun which is marked for number. Consequently, they specify whether the number of

individuals that match the description is singular or plural. Nevertheless, the exhaustiveness condition behaves in the exact same way. For example, in (25), *Howard and Paul* are understood to be the only people that left. This is a consequence of the fact that *the ones that left* is a universally quantified set, encompassing all of its members.

(25) The **ones** that left were Howard and Paul [specificational *NP be NP*]

From this Russellian analysis, the exhaustiveness of *it*-clefts can therefore be traced to the semantic contribution of the discontinuous constituent (involving *it* and the extraposed relative clause), which is a definite description. However, there are other cleft constructions, such as *there*-clefts, which are not exhaustive. The function of sentences such as (26) seems to be to acknowledge that the speaker is unsure of all of the members of the described set; that is, they don't really know what the set constitutes.

(26) Well, there's John and Mary that are available to work Saturday [...and I think maybe Bob too...I'll ask around]

Nevertheless, these too correspond to a subset of noncleft specificational sentences. While exhaustiveness is the necessary outcome of the standard use of definite or definite-like descriptions in specificational sentences, indefinite noun phrases do not involve universal quantification. Consequently, specificational sentences with indefinite subjects are not usually exhaustive, since the focal element only ever identifies a single member of the (potentially plural) set (see §3.3.4). Take the examples in (27).

(27) A scientist that is currently working on the new drug is Dr. John Matthews... Another scientist working on the drug is Dr Andrew Perry

In this example, two indefinite specificational sentences follow on from one another, citing different individuals as corresponding to the same set. *Contra* Declerck (1988), exhaustiveness is therefore not inherent to the act of specification, but is a consequence

of the semantic contribution of the definite article and the (functionally similar) constituents *what* and *it* in clefts.

In what follows, I compare my semantic explanation for the exhaustiveness of *it*-clefts to the syntax-based account proposed by É. Kiss (1998). As I explained in §3.1, É. Kiss (1998) argues in favour of an expletive analysis of *it*-clefts; that is, she assumes that the cleft clause is directly predicated of the clefted constituent and that the cleft pronoun is a dummy element. Here, I show that without recognizing the important semantic contribution of the initial *it*, É. Kiss cannot satisfactorily account for the property of exhaustiveness.

É. Kiss (1998: 251) claims that "exhaustive identification is a function of structural focus". She suggests that, cross-linguistically, this "identificational focus" occupies a designated syntactic position as the specifier of a functional projection (the focus phrase) and goes on to claim that the clefted constituent in *it*-clefts "is the realization of identificational focus in English" (É. Kiss 1998: 258). As a result, É. Kiss provides a syntactic analysis of *it*-clefts in which the clefted constituent is moved from within the cleft clause, through spec-CP and into the specifier slot of the focus phrase, as shown in (28)[11].

(28) $[_{IP}$ It $[_{I'}$ was$_k$ $[_{FP}$ HOWARD$_i$ $[_{F'}$ t_k $[_{CP}$ $[t_i$ $[_{C'}$ that $[_{IP}$ t_i left]]]]]]]]

After moving into this scope position, the clefted constituent performs two important roles. Syntactically, the clefted constituent functions as an operator, which marks "the sentence part following it and c-commanded by it as the scope of exhaustive identification" (É. Kiss 1998: 253). Semantically, the clefted XP expresses the complete set of elements for which the predicate in the cleft clause holds.

There are some important problems with É. Kiss' (1998) account. For one thing, her analysis cannot account for the full range of data. É. Kiss assumes that exhaustiveness is dependent upon a syntactic structure which in English is particular to

[11] Alternatively, the clefted constituent can be base-generated in spec-FP and then coindexed with the corresponding *wh*-pronoun in the cleft clause (É. Kiss 1998: 259). This allows É. Kiss to account for all of the *it*-cleft's agreement patterns (see §4.3).

the *it*-cleft configuration. However, we have seen above that this property is also found in other specificational copular constructions. It is not at all clear whether É. Kiss' identificational focus structure could be accommodated into the *NP be NP* syntactic configuration of examples such as (22) and (25) above. *There*-clefts also pose a problem for É. Kiss' claim that exhaustive identification is tied to cleft structure. As shown in (26) above, *there*-clefts are not exhaustive. Since these sentences presumably have the same structure as *it*-clefts, this is not expected under É. Kiss' analysis.

In addition, É. Kiss' (1998) analysis is highly stipulative and suffers from circularity. É. Kiss (1998: 245) assumes that the clefted constituent functions both as a syntactic operator (expressing exhaustive identification) which binds the variable left behind in the cleft clause and as the value for this variable. In other words, she assumes that the same constituent that expresses an exhaustive set of elements is also the constituent that causes exhaustiveness in *it*-clefts. This does not make a great deal of sense to me and, since her analysis is specific to the cleft configuration, it cannot be supported by more general patterns of correspondence in the language.

As a result, despite É. Kiss' (1998) attempts, expletive accounts cannot successfully account for the exhaustiveness in *it*-clefts. This is only possible once we recognize the important semantic contribution of the constituent *it*. Once we take this step, we can explain the full range of data in an analysis that is supported by inheritance from definite noun phrases.

4.1.2.4 Contrast

Along with exhaustiveness, it is often noted that the postcopular element of a cleft sentence is contrastive. For example, the positive cleft in (29) contrasts *Howard* with the set of people at the party who didn't leave early. This sense of contrast is even more pronounced in negative clefts. For example, in (30) *Howard* is contrasted with the individual that did leave early.

(29) It was Howard that left the party early

(30) It wasn't Howard that left the party early

I suggest that the contrastive nature of cleft sentences is a consequence of the exhaustive and existential presuppositions of definite noun phases. Exhaustively listing the complete set of individuals that left early in (29) invites a sharp contrast between those people that left early and those people that didn't. Likewise, in (30) the existential presupposition that there is a set of people that did leave early invites a contrast between *Howard*, who is excluded from this set, and the unknown individual or individuals that together make it up. This sense of contrast is heightened, since *Howard* is being contrasted with other individuals that attended the party rather than all other people that exist.[12] In these examples, the set of possible referents that could potentially make up the described set is restricted by information given in the relative clause.

A contrastive interpretation is also possible for noncleft specificational sentences. For example, depending upon the context, (31) can be used to contrast *Dr. Amy Schrute* with other *non-psychologists working at the hospital* or with other *psychologists that don't work at the hospital*. In the former interpretation, the restrictive relative clause provides the "background set" of hospital workers, with the head noun providing the characteristic that distinguishes the referent from all other members of this background set. In the latter interpretation, the head noun *psychologist* provides the background set, with the relative clause telling us what separates the referent from all other psychologists. However, if the restrictive relative clause is omitted, the head noun must be interpreted as containing the "distinguishing" information. For instance, example (32) can only be used to contrast *Dr. Amy Schrute* with other *non-psychologists (that work at the hospital)*.[13] It cannot be used to contrast the referent with other psychologists.

(31) The psychologist that works at the hospital is Dr. Amy Schrute

(32) The psychologist is Dr. Amy Schrute

---

[12] Declerck (1988: 25) also notes that the sense of contrast is stronger when the set from which the referent is selected is restricted. He provides the example given here as (i).

(i) We don't know which of them is the murderer, but it certainly isn't Tom (Declerck 1988: 25)

[13] In this section, I use parentheses to indicate information which must be recovered from the context.

This provides a further explanation as to why the restrictive relative clauses of cleft sentences cannot be omitted (see examples (10) and (11) above and the surrounding discussion). Since the head noun of the precopular phrase in clefts is underspecified, it is the context that tells us what it denotes. For example, in a discourse about psychologists, the precopular phrase in (33) is understood to mean *the one (out of the set of psychologists) that works at the hospital*. The referent *Dr. Amy Schrute* can be contrasted with other *psychologists that don't work at the hospital* but not with *non-psychologists that work at the hospital*. In cleft sentences then, the head noun always denotes the "background set" from which the referent is selected, with the restrictive relative clause providing the information that distinguishes the referent from the other members of this background set. Since, as shown above, only the background set can be left unexpressed, it follows that the restrictive relative clause in cleft sentences cannot be omitted. For example, in (34) the noun *one* does not provide any information with which to restrict the referent from other members of the background set of *people working at the hospital*.

(33) The one that works at the hospital is Dr. Amy Schrute

(34) *The one is Dr Amy Schrute

In order to capture the two possible meanings of (31) above, cleft sentences must alter the information in the relative clause. For example, in a discourse about the hospital, (35) can be used to contrast *Dr. Amy Schrute* with *non-psychologists (that work at the hospital)*.

(35) It's Dr. Amy Schrute that's the psychologist

As well as involving existential and exhaustiveness presuppositions, contrast in cleft sentences is therefore also a consequence of the fact that definite descriptions denote restricted sets. Containing both a head noun and further restrictive information, definite noun phrases allow for two possible interpretations depending upon the context:

either the head noun denotes the background set which is further restricted, or the background set is provided by the (potentially unexpressed) modifying information, with the head noun telling us what distinguishes the postcopular referent from the other members of this set. In essence then, definite noun phrases denote restricted sets by way of restricting a slightly more general set. In specificational sentences (where a referent is identified as matching the description of the restricted set) it follows that this referent is contrasted with other members of the background set.

4.1.2.5 Givenness

Finally, it is well known that the sentence-final clause in cleft sentences is made up of given, or known, information. For instance, in the following exchange, the postcopular element expresses the new information, that of the referent *Paul*, whereas the relative clause *stole the money* is discourse-old information.

(36) A: Which of the students **took it**? B: It was Paul that **stole the money**

By interpreting the relative clause as corresponding to the restrictive set information of a definite description, this fact becomes unsurprising. As we have seen, definite noun phrases involve existential commitment. If the interlocutors assume that the described set exists, it follows that they are likely to be familiar with that set. Consequently, the use of definite descriptions is associated with given or shared information (see also §3.3.4). The example above therefore represents a "prototypical" *it*-cleft, in which the unknown member (in this case *Paul*) of a known set of *students that stole the money* is identified. However, as I go on to explain in §5.6, not all *it*-clefts are associated with given information. Such examples require an independent explanation.

4.1.2.6 Summary and interim conclusions

Throughout this section, I have provided additional support for the analysis proposed in §4.1.1, in which the constituent *it* and the cleft clause function together as a discontinuous definite description. I have shown that many of the *it*-cleft's pragmatic

properties fall out straightforwardly from this analysis and can be explained either as a product of specificational meaning (namely focus), or as a resulting from the inherent semantics of definite noun phrases (including existentiality, exhaustiveness, contrast and givenness).

I conclude that the analysis put forward in this thesis has much more explanatory power than the expletive accounts outlined in §3.1. For proponents of the expletive approach, the constituent *it* is a semantically inert element. Without recognizing that the initial *it* performs a definite-like quantificational role, expletive accounts cannot satisfactorily explain where these pragmatic properties come from. For example, while É. Kiss (1998) invokes construction-specific, stipulative mechanisms to account for the exhaustiveness in *it*-clefts, I have shown that these pragmatic properties are motivated by inheritance from the wider specificational construction and are supported by more general patterns of correspondence in the language system.

4.1.3 Explaining the behaviour of the cleft clause

In addition to explaining the *it*-cleft's pragmatic properties, the analysis outlined in this chapter also helps us to understand the construction's structural properties. I have argued in favour of a discontinuous constituent analysis of *it*-clefts in which the cleft clause and the constituent *it* form a semantic unit, acting as a discontinuous definite description. In my own version of this type of approach, the cleft *it* is equivalent to *the one* in *th*-clefts. Consequently, this element is not classified as a complete and referential NP, but as performing definite-like quantification and acting as the underspecified head noun of the description. From this, it follows that the cleft clause is a restrictive relative, modifying the constituent *it*.

In this section, I provide further evidence in support of this analysis. I show how once we view the cleft clause as modifying the constituent *it*, its behaviour is consistent with a restrictive relative clause analysis. In this way, my analysis of *it*-clefts has an important advantage over expletive accounts, which view the cleft clause as being in some way related to the complement of *be*. On this type of analysis, the behaviour of the

cleft clause is unpredictable and has to be stipulated as idiosyncratic to the cleft construction.

As I explained in §3.1, scholars who adopt the expletive analysis recognize that the sentence-final clause in *it*-clefts looks and behaves in some ways like a relative clause. For some of these authors, there is only a superficial similarity between these structures. For example, Delahunty (1982), Rochemont (1986) and Heggie (1988) suggest that the occurrence of relative pronouns in the cleft clause of *it*-clefts with NP foci, such as *who*, is the result of analogy with relative clauses. Others, such as Chomsky (1977) and Williams (1980), claim that the cleft clause has the same internal structure as a restrictive relative but note that the cleft clause does not restrictively modify its immediate antecedent. Instead, as Heggie (1988) notes, the external relationship between the cleft clause and the clefted constituent has more in common with the behaviour of nonrestrictive relatives.

Still other proponents of the expletive analysis have overtly classified the cleft clause as a type of relative clause. For example, Delin (1989) suggests that when the postcopular element is nominal, as in (37), the cleft clause can be given a restrictive relative analysis. However, as I noted in §1.1, *it*-clefts can occur with a range of elements in the postcopular position. Since restrictive relative clauses cannot modify non-nominal antecedents, Delin claims that in these examples, such as the prepositional phrase focus *it*-cleft in (38), the sentence-final clause should be analysed as a sentential complement. For Lambrecht (2001), on the other hand, the cleft clause is a type of nonrestrictive relative clause. This allows him a unified analysis of all kinds of *it*-cleft, since nonrestrictive relatives can have non-nominal antecedents, shown in (39).

(37) It was **the vicar** who had him shot

(38) It was **with a knife** that he cut it

(39) She quickly tucked the key **under the mattress**, which was her favourite hiding place, and returned to her book.

According to Lambrecht (2001: 473), the cleft clause is predicated of the focal element but it does not perform a modifying function; this property is shared with other nonrestrictive relatives.

Both Delin (1989) and Lambrecht (2001) therefore attempt to integrate the cleft clause (at least for *it*-clefts with nominal focal elements) into a unified analysis of relative clauses. While this is clearly advantageous, there is an important problem with analysing the clefted constituent as the antecedent to the relative clause: the relationship between the cleft clause and its immediate antecedent appears to be somewhat different from that found with relative clauses in other constructions. Consequently, on an expletive account, the relative clause in *it*-clefts does not fit neatly into either the restrictive or nonrestrictive category. As Lambrecht (2001: 468) acknowledges, "Its category membership is sometimes debatable or unclear".

On the one hand, the cleft clause shares a number of properties with nonrestrictive relatives (if we assume that it is relative to the complement of *be*). For instance, the immediate antecedent to the cleft clause can be a pronoun or proper name, as in (40). Nonrestrictive relative clauses, which provide additional information about their antecedents, can also attach to these elements, as in (41). In contrast, restrictive relative clauses modify and restrict the "type specification" of the antecedent noun (Langacker 1991: 432), shown in (42). Since pronouns and proper names are full noun phrases, they cannot be further modified by restrictive relatives (see example (43))[14].

(40) It's Howard that plays the bassoon

(41) Howard, who's a specialist in snakes, cured him

(42) I gave it to the man that wears a green jacket

(43) *I gave it to Howard that wears a green jacket

[14] As Davidse (2000: 1111) points out, a restrictive relative clause can modify a proper noun only if it is preceded be a determiner, as in (i). Here, *John* is used not to refer to a fully specified entity, but as a common noun denoting a general type "person with the name *John*".

(i) I've just met the John you used to go out with (Davidse 2000: 1111)

As Delin (1989: 60) acknowledges, this poses a problem for her claim that NP-focus *it*-clefts contain restrictive relative clauses.

Cleft clauses also appear to pattern with nonrestrictive relatives in terms of their prosodic structure. Restrictive relative clauses and their antecedent nouns form a constituent, and therefore an intonation unit, within the noun phrase, with nuclear stress falling on the tone-final element (see example (44)). However, attaching as they do to full noun phrases, nonrestrictive relatives occur as a separate constituent. Like these nonrestrictive relatives, the cleft clause does not form an intonation unit with its antecedent, shown in (45) (see Halliday 1967: 237; Davidse 2000: 1103).

(44) [The doctor who phòned me] lives there

(45) It was [the dòctor] [who phóned me]

However, the relative clause in *it*-clefts has a number of other characteristics that are typical of restrictive, rather than nonrestrictive relative clauses. For instance, while nonrestrictive relatives are separated from the rest of the sentence by commas in writing or pauses in speech, this property is not shared by restrictive relatives or the cleft clause. The sentence-final clause in *it*-clefts also occurs with the same range of introductory elements as restrictive relatives. For example, the cleft clause is commonly introduced by *that* and can also occur without an overt relative pronoun. While restrictive relatives commonly occur with *that* or zero, both options are strongly resisted by nonrestrictive relative clauses, which typically require an overt *wh-* relative pronoun.

(46) It was Howard that/Ø I used to work with [*it*-cleft]

(47) The guy that/Ø I used to work with lives there [restrictive]

(48) *Howard, that/Ø I used to work with, lives there [nonrestrictive]

This evidence poses a problem for the analysis of the cleft clause as a nonrestrictive relative. However, Lambrecht's (2001) "nonrestrictive" analysis of the cleft clause is dependent more on the lack of similarity to restrictive relatives rather than

on perceived correspondences with other nonrestrictive relative clauses. For example, while restrictive relative clauses only permit zero realization of the relative marker when the antecedent acts as an object complement of the proposition in the relative clause, *it*-clefts, such as (50), can occur without *that* even when the relative element is the subject (see also Sornicola 1988: 347).

(49) *The man paid for that is over there [restrictive relative clause]

(50) It was your husband paid for that [*it*-cleft]

(Lambrecht 2001: 470; Delahunty 1982: 52)

Rather than analysing the example in (50) as involving "(substandard) complementizer drop", Lambrecht (2001: 470-471) takes this as evidence that, when present, the complementizer *that* in *it*-clefts is semantically inert; "we could say with equal if not greater justification, that this sentence represents a canonical sentence (*Your husband paid for that*)". From this, it follows that the cleft clause is directly predicated of the clefted constituent. While Lambrecht (2001: 473) argues that the cleft clause is not a restrictive relative, he notes that "the fundamental property of all RCs [is] that they are PREDICATES" (emphasis original). On this basis, Lambrecht categorizes the cleft clause as a "kind of" nonrestrictive relative.

The attempts of Delin (1989) and Lambrecht (2001) to categorize the sentence-final clause in *it*-clefts as a type of relative clause are therefore highly problematic. As we have seen, if we assume that the clefted constituent is the antecedent to the cleft clause, then this clause exhibits properties of both restrictive and nonrestrictive relatives. Consequently, with an expletive analysis in place, the cleft clause does not conform to the general principles of the grammar and requires an idiosyncratic account. As Huddleston (1984: 462) comments, this analysis, in which the clefted constituent is perceived as the antecedent to the cleft clause, "is very largely ad hoc – the relative clause is of a kind that is sui generis, unique to this construction".

However, we can rescue the relative clause analysis of the cleft clause by treating the *it*-cleft as a fully-fledged specificational construction. Now, the antecedent to the

relative clause is shown to be not the complement of *be* (the clefted constituent), but the initial *it*, which quantifies the description as well as acting as head noun. Invoking this analysis, the cleft clause can be given a reasonably consistent analysis as a restrictive relative clause. This analysis has two advantages: it helps cash out the treatment of the initial *it* and the relative clause as a definite description and it maximizes generality and consistency within the grammar.

Regardless of the syntactic category of the clefted constituent, in this account the antecedent to the cleft clause is the nominal *it*. As a result, there is no need to establish a separate function of "sentential complement" for *it*-clefts with non-NP foci, *contra* Delin (1989). We can see then that example (51) below corresponds to (52), in which the relative clause modifies the abstract, adverbial noun *way*.

(51) [It] was with a knife [that he cut it]

(52) [The way he cut it] was with a knife

The constituent *it* normally functions as a complete referential pronoun, and so would not normally be the expected antecedent of a restrictive relative. However, by analysing the cleft *it* as **part** of a definite-like description, equivalent to *the one* or *the thing*, it is possible for the relative clause to modify this element restrictively. Although this function for the initial *it* is therefore specific to the *it*-cleft construction, I provide additional evidence for this analysis of the cleft *it* in §7.1.

Like restrictive relatives, the cleft clause forms a constituent and therefore an intonation unit with the unstressed *it*, shown in (53). Huddleston (1984: 461) notes that an analysis in which the relative clause modifies *it* rather than the complement of *be* clearly accounts for the separate readings of examples such as (53) and (54) below. In (53), the relative clause restrictively modifies *it*; in (54), *the doctor* is the antecedent.

(53) A: Who phoned? B: [It] was the dòctor [who phóned me]

(54) A: Who was at the door? B: It was [the doctor who phòned me]

By treating the constituent *it* as the antecedent of the relative clause, it is also possible to explain why *it*-clefts can accommodate the omission of the relative pronoun, even when the relative element is the subject of this embedded clause. While (55) creates an unintended "garden-path" reading, in which *the man ∅paid for that* can be mistakenly understood as a complete subject-predicate sentence, this ambiguity does not arise for the *it*-cleft in (56), since the antecedent of the relative clause is not the complement of *be*.

(55) *The man paid for that is over there [restrictive relative clause]

(56) It was your husband paid for that [it-cleft]

Consequently, and in comparison to the expletive accounts, by treating the constituent *it* and the cleft clause as forming a definite description, the sentence-final clause in *it*-clefts can be unequivocally classified as a restrictive relative, and does not have to be treated separately as idiosyncratic to this construction. With this analysis in place, the "unusual" behaviour of the cleft clause becomes predictable from what we know about how restrictive relative clauses operate in the rest of the grammar.

4.1.4 Summary and interim conclusions

Throughout §4.1, I have shown that an analysis which views the *it*-cleft primarily as a specificational copular construction has greater explanatory force than an expletive account. I began by arguing that *it*-clefts, like the specificational *NP be NP* sentences analysed in §3.3, involve a special type of nominal predication relation. This led me to advocate a discontinuous constituent analysis of *it*-clefts in which the initial *it* and the cleft clause form a definite-like description. I have shown that this unified analysis of *it*-clefts and other specificational copular constructions allows the maximum number of the *it*-cleft's formal and functional properties to be inherited from the wider specificational construction and from more general patterns of correspondence in the language. This account explains both why these properties occur and what roles they perform in the creation of specificational meaning.

4.2 A Comparison with other extraposition accounts

In §4.1, I explained how a semantic account of specificational copular sentences calls for a (non-derivational) extraposition analysis of *it*-clefts in which the constituent *it* and the relative clause function as a discontinuous definite description. I have shown that this analysis has a number of advantages over the expletive accounts and is able to provide motivation and an explanation for many of the *it*-cleft's seemingly idiosyncratic properties. In this section, I review a selection of alternative extraposition accounts and explain how my own particular version of this approach draws on, and improves on, the existing literature.

4.2.1 Jespersen (1927)

My analysis of cleft sentences shares much with Jespersen's (1927) original "transposition theory".[15] Prior to rejecting these ideas in favour of an expletive account, Jespersen analysed the sentence-final clause in *it*-clefts as a restrictive relative. Since this clause occurs immediately after a phrasal constituent that cannot be further modified, he suggests that the relative clause restricts the initial pronoun *it* rather than the focal element. In this way, Jespersen captures the close relationship between *it*-clefts, such as (57), and other types of specificational sentence, such as the reverse specificational example given here as (58).

(57) It was the Colonel I was looking for

(58) The colonel was the man I was looking for (examples from Jespersen 1927: 88)

This early account therefore foreshadows the type of 'discontinuous constituent' analysis advocated in this thesis (see also §4.2.7, §4.2.8 and §4.2.9 below).

---

[15] Although often cited, Jespersen (1927) was not the first to propose an extraposition analysis of *it*-clefts. For instance, Fowler and Fowler (1908) suggest that the cleft pronoun and the cleft clause function together as a subject NP: "the 'that' clause, supplemented or introduced by 'it', gives us the subject of a predication", while "the predication answers an imaginary question recorded distinctly in the relative". On this analysis, the *it*-cleft *It is money that I want* is understood in relation to the exchange "'What do you want?' 'It (the thing) that I want is money'".

4.2.2 Akmajian (1970)

The transformational accounts of the 1970s share Jespersen's insight that *it*-clefts should be examined in relation to other specificational copular constructions. However, not all of these analyses assume, as Jespersen (1927) and I do, that the sentence-final clause restrictively modifies the constituent *it*. For example, Akmajian (1970) suggests that *it*-clefts are syntactically derived from "headless" pseudocleft sentences. Using the terminology adopted in this thesis, Akmajian claims that *it*-clefts are derived from *wh*-clefts, such as (59a) and (60a), but not from *th*-clefts, such as (61). Although *wh*-clefts introduced by *who* are regarded by many as ungrammatical, Akmajian (1970: 164) includes these examples, such as (60a), as suitable "sources" for *it*-clefts, noting that for some speakers, these examples "*obligatorily* become cleft sentences" (italic original).

(59) a) What John bought was a car
b) It was a car that John bought (Akmajian 1970: 150)

(60) a) ??Who chose Nixon was Agnew
b) It was Nixon who chose Agnew (Akmajian 1970: 149)

(61) The one who Nixon chose was Agnew

On this account, *it*-clefts are derived from *wh*-clefts by a rule particular to cleft sentences: the "Cleft Extraposition Rule", which may require the replacement of the *wh*-word by *that*, as in (59) above. Akmajian (1970: 150) claims that, after the extraposition of the relative clause, "the element IT is left in subject position". Because Akmajian does not include the "head" of the relative clause in the derivation of *it*-clefts, he fails to see the link between *the one* in (61) and *it* in (60b).[16] For Akmajian (1970) then, the constituent *it* is an expletive element and is not restricted by the sentence-final clause.

[16] Akmajian claims that *it*-clefts cannot be derived from *th*-clefts because they often do not have the same meaning. For example, Akmajian (1970: 161-162) notes that the headed pseudocleft in (i) tells us "that a certain *place* is located in the garden" while the headless pseudocleft in (ii) and the *it*-cleft in (iii) tell us "that *John* was located in a certain place". Akmajian seems to be claiming here that *th*-clefts have a predicational meaning while *it*-clefts and *wh*-clefts are specificational. However, Bolinger (1972) successfully argues against this claim, showing that both *th*-clefts and *wh*-clefts allow both predicational and specificational (or, using his terminology, "equational") readings.

To my mind, Akmajian's (1970) analysis misses an important generalization. In §4.1.1, I analysed the *wh*-clause in *wh*-clefts as a fused relative construction rather than as a "headless" relative, claiming that the initial *what* in *wh*-clefts performs the same role as *the thing that* in *th*-clefts. In other words, the *wh*-clause functions as a complete definite noun phrase – this is what allows the *wh*-cleft configuration to acquire a specificational meaning. On my analysis, it follows that if this *wh*-clause is extraposed then the resulting construction is a 'right-dislocated pseudocleft' rather than an *it*-cleft.

4.2.3 Gundel (1977)

Gundel (1977) agrees, providing an analysis in which *it*-clefts are derived from right-dislocated structures. In right-dislocated pseudoclefts, the initial fused relative of the *wh*-cleft is extraposed, with *it* acting as a pronominal reference to the right-dislocated constituent, shown in (62b). Gundel claims that *it*-clefts are reduced forms of right-dislocated pseudocleft, derived by an optional, construction-specific "Variable Head Deletion" rule. This rule deletes the head element of the compound pronoun *what*, whilst simultaneously deleting the intervening sentence boundary between the sentence-final clause and the postcopular constituent. The result, she claims, is an *it*-cleft with a sentence-final restrictive relative clause structure and a referential subject (*it*) which acts as a pronominal copy for the dislocated constituent.

(62) a) What you heard was an explosion
b) It was an explosion, what you heard
c) It was an explosion that you heard (examples from Gundel 1977: 543)

However, an unfavourable consequence of Gundel's account is that right-dislocated pseudoclefts, which appear less natural than their *it*-cleft counterparts, are analysed as having a more basic structure. Furthermore, as Gundel (1977: 557) acknowledges, the Variable Head Deletion rule "has no independent motivation in

---

(i) The place where I found John was in the garden
(ii) Where I found John was in the garden
(iii) It was in the garden that I found John (examples from Akmajian 1970: 161-162)

English, and the fact that it can apply only in ID [specificational] structures is admittedly suspicious". Gundel's analysis is therefore *ad hoc*: it involves a major constructional change from a complete noun phrase into a restrictive relative without explaining why a restrictive relative clause structure should be able to occur without its normal modifying function. Consequently, Gundel's analysis leaves too many questions unanswered and provides a structure for the *it*-cleft that increases, rather than reduces, the number of idiosyncratic properties that the construction exhibits.

4.2.4 Wirth (1978)

For Wirth (1978), both *it*-clefts and pseudoclefts are derived from the same source: from copular sentences whose subjects contain a relative clause with a pronominal head. *It*-clefts are derived via an extraposition rule and a "restricted nominal reduction" rule which leads to the realization of the pronominal head as *it*. Wirth's account captures the idea that *it* and the definite article share the same function, claiming that *it* is a "syntactic variant" of *the*, occurring when a head noun is not stipulated. However, it is not clear what semantic value Wirth attributes to the constituent *it*. Furthermore, like Akmajian (1970), Wirth suggests that *wh*-clefts involve a "headless" relative structure. Consequently, she too fails to provide a unified analysis that recognizes the role of definite quantification in all of these specificational constructions.

4.2.5 The problem with transformational accounts

Despite their differences, the accounts of Akmajian (1970), Gundel (1977) and Wirth (1978) all share an additional, common problem which stems from their adoption of transformational theories of grammar. These authors all assume that *it*-clefts are derived from alternative copular structures, including *wh*-clefts, right-dislocated pseudoclefts and copular sentences with subjects containing modified pronominal heads. On a transformational account, it is assumed that the source structure and all of the derived structures will exhibit the same behaviour. However, *it*-clefts differ in some important ways from *wh*-clefts (right-dislocated or otherwise).

For example, as I go on to explain in §5.3, *wh*-clefts and *it*-clefts allow a different range of focal elements to occur in the postcopular position. In some ways, the range of possible *it*-cleft foci is more restricted. For instance, adjectival foci are much more acceptable in *wh*-clefts (shown in (63)). In order to account for this, Akmajian (1970) and Wirth (1978) provide restrictions as to which types of copular sentence the transformational rule of extraposition can apply to. Akmajian (1970) suggests that the Cleft Extraposition Rule only applies to pseudoclefts with NP or PP foci, while Wirth (1978) argues that extraposition is only possible over postcopular noun phrases. However, in truth, the range of possible *it*-cleft foci is much greater than this and cannot be divided into major phrasal categories. For instance, *it*-clefts can sometimes accommodate adjectival foci, but only in certain circumstances, shown in (64).

(63) a) What John is is **tall**
b) *It is **tall** that John is

(64) It's not **sick** that he was but tired (É. Kiss 1998: 262)

Even more problematic for these transformational accounts is the fact that some *it*-clefts have no possible pseudocleft source. For example, certain prepositional phrases will only be permitted as the focal element of an *it*-cleft, shown in (65a) and (65b). In order to accommodate these examples into her transformational account, Gundel (1977) suggests that they are instead derived from other cleft sentences, such as (65c), by an optional rule which copies the preposition into the focus and deletes the original.

(65) a) It was with George that Mary eloped (Gundel 1977: 550)
b) *(The one) that Mary eloped was with George
c) It was George that Mary eloped with (Gundel 1977: 550)

However, Preposition Copying cannot account for all *it*-cleft examples without a pseudocleft source. For example, Gundel suggests that (66a) cannot be derived from the *wh*-cleft in (66b). However, the *it*-cleft in (66c) is at least an equally unacceptable

source.[17]

(66) a) It is with great pride that I accept this nomination (Gundel 1977: 548)
b) *How I accept this nomination is with great pride (Gundel 1977: 548)
c) *It is great pride that I accept this nomination with

So how does this compare with the constructional approach which I propose? Throughout this chapter, my main concern has been to maximize correspondence between the *it*-cleft and other specificational copular constructions. As a result, the discussion has so far centred on *it*-clefts with NP foci, since these examples show the greatest similarity to the specificational *NP be NP* sentences discussed in chapter 3. However, a key property of the construction grammar framework is that while the motivation for each construction should be maximized, idiosyncrasies are both tolerated and expected (see §2.5). Therefore, the fact that the *it*-cleft permits a construction-specific range of foci, which is neither shared by other specificational constructions nor comprised of major syntactic categories, is not problematic for this analysis. Consequently, and in contrast to a transformational account, a constructional analysis allows us to observe true generalizations without requiring us to provide a series of *ad hoc*, unmotivated and complex rules in order to explain away every construction-specific property that does not (intuitively) seem to be the product of a more general pattern. I return to this issue, and to non-NP focus *it*-clefts in chapter 5.

---

[17] Emonds (1976) puts forward an alternative account which proposes to handle examples which have no possible pseudocleft source, such as (65a) repeated here as (iii). Emonds' analysis is really a mixture of expletive and extraposition approaches. He assumes that *it*-clefts are derived from a source structure in which the clefted constituent is base-generated inside the cleft clause, shown in (i). Emonds claims that the clefted constituent (NP or PP) is moved out of the cleft clause by a transformation called "focus placement" before the structure undergoes Akmajian's (1970) "cleft extraposition".

(i) [that Mary eloped with George] was
(ii) [that Mary eloped] was with George
(iii) It was with George that Mary eloped

However, Emonds' analysis is highly suspect. For one thing, his analysis predicts that *it*-clefts permit only NP or PP foci. Furthermore, as É. Kiss (1998: 258) argues, Emonds' set of derivations "is highly stipulative; the initial structure is unlikely, and the rightward movement rule focus placement is not independently motivated".

So far then, I have shown that the transformational accounts of Akmajian (1970), Gundel (1977) and Wirth (1978) do not compare favourably with the extraposition analysis argued for in this thesis. We can now look at some extraposition proposals which are more similar to my own. All of the following analyses claim, as I do, that the initial *it* and the cleft clause form a semantic unit.

4.2.6 Bolinger (1972)

Bolinger (1972) proposes a somewhat different transformational account to Akmajian (1970), Gundel (1977) and Wirth (1978). Although he suggests that analytic compound relatives, such as (67), "provide an ideal source for cleft sentences, since they spell out the two elements that are divided between clauses", Bolinger (1972: 110) does not provide an analysis whereby *it*-clefts are derived from these structures. Instead, Bolinger argues that the same transformational rule creates the "inverted" structure in (68) and the *it*-cleft in (69). This rule involves the "inversion" of the restrictive relative clause while the nominal head remains in initial position. Bolinger claims that while "inversion" (or extraposition-from-NP) is optional for (67), it applies obligatorily in *it*-clefts. Consequently, the transformational source of the example given in (69) is not an acceptable sentence, as shown in (70).

(67)	**That which** he stole was money	[analytic compound relative]
(68)	**That** was money **which (that)** he stole	['inverted' compound relative]
(69)	**It** was money **that** he stole	[*it*-cleft]
(70)	***It that** he stole was money	[*it*-cleft source]

(examples from Bolinger 1972: 109)

Bolinger's (1972) account exhibits a number of similarities with both Jespersen's (1927) proposal and the analysis provided in this thesis. *It*-clefts are analysed as containing an extraposed restrictive relative clause which modifies the initial *it*. Bolinger's analysis therefore contrasts with the proposals of Akmajian (1970) and Gundel (1977) since it suggests that the constituent *it* is important to the semantics of the

construction and is not simply an expletive element or a pronominal copy which is inserted as a consequence of extraposition.

However, Bolinger's (1972) "obligatory inversion" for *it*-clefts raises an important question. If a transformational rule is obligatory, how can we be sure that speakers actually have the underlying source structure? In the case of the *it*-cleft, there seems to be no motivation for this kind of transformational analysis. A constructional approach such as the one advocated in this thesis assumes that *it*-clefts are not derived from any other structure. Since construction grammar is a monostratal theory, underlying structures are not posited.

4.2.7 Percus (1997)

Recently, a number of extraposition analyses have been put forward which explicitly argue, as I do, that the constituent *it* and the relative clause of *it*-clefts function together as a discontinuous definite description. For example, Percus (1997) argues that the sentence-final clause in *it*-clefts is an extraposed restrictive relative clause modifying the initial *it*. As in my analysis, the constituent *it* provides the functions of the definite article as well as acting as the head noun of the description. Percus notes that this type of analysis explains why *it*-clefts share a number of properties with definite descriptions, including existential and exhaustiveness presuppositions.

However, Percus' (1997) account differs from my own analysis in some important respects. While I claim that the cleft *it* and the relative clause function together like a discontinuous definite description, inheriting properties from specificational sentences with definite NPs, Percus argues that *it*-clefts are actually derived from an underlying structure containing a definite description. For example, Percus (1997: 338) assumes that the *it*-cleft in (71a) has the underlying structure given here as (71b). This is subject to extraposition-from-NP, shown in (71c), before the definite determiner and null head are realized as *it* at morphological spell-out, given in (71d).

(71) a) It is $[\text{JOHN}]_F$ that Mary saw.

b) $[_{IP}\ [_{DP}$ the 0 $[_{CP}$ OP_i that Mary saw $t_i]]_j$ $[_{VP}$ t_j is John]]

c) $[[_{IP}\ [_{DP}$ the 0 $t_k]_j$ $[_{VP}$ t_j is John]] $[_{CP}$ OP_i that Mary saw $t_i]_k]$

⇓ ⇓ ⇓ ⇓ ⇓ ⇓

d) **It** is John that Mary saw

Spell-Out: $[_{DP}$ the 0 $t_k] \Rightarrow$ It (Percus 1997: 338)

In addition to the theory of grammar that Percus adopts, his understanding of *it*-cleft semantics differs crucially from my own. Percus (1997: 346) argues that specificational sentences, including *it*-clefts, are semantically equative, relating two arguments of the same type. Percus notes, however, that in specificational sentences, one of these arguments must refer to an "unknown" individual. Consequently, although Percus does not mention Donnellan's (1966) referential/attributive distinction, his analysis is similar to that of Declerck (1988), and is therefore subject to the same fundamental problems (see §3.2.3.2). Percus claims that predicational and equative structures contain two different types of copular verb. While the former relates an individual and a predicate and does not allow inversion, the latter can be inverted, "selecting two arguments of the type of individuals" with the restriction that "one of its syntactic arguments be a D_u ['unknown' description]" (Percus 1997: 346).

4.2.8 Hedberg (1990, 2000)

Hedberg (1990, 2000) also advocates a discontinuous constituent analysis of *it*-clefts, formulating two slightly different versions of this approach. I begin by describing Hedberg (2000). For Hedberg (2000), the constituent *it* and the sentence-final clause of the *it*-cleft function pragmatically and semantically like a definite description. Hedberg (2000: 894) proposes that definite noun phrases are comprised of two parts: "an INDEXICAL component, which is expressed by the determiner head and determines the relation of the referent to the context, and a DESCRIPTIVE component, which is expressed by its nominal complement and describes the referent" (emphasis original). For Hedberg (2000: 898), "*it* is analyzed as an allomorph of *the*", acting as the definite article head

(the indexical component) of the determiner phrase, while the sentence-final clause is said to function as the descriptive component. This is shown in (72) and (73), where the descriptive component is highlighted in bold.

(72) [$_{DP}$ [$_{D}$ the] **[$_{NP}$ [$_{NP}$ dog] [$_{PP}$ next door]]]**

(73) [$_{DP}$ [$_{D}$ it]…**[$_{CP}$ who won]]** (examples from Hedberg 2000: 898)

Consequently, Hedberg's analysis differs both from that proposed by Percus (1997) and from the account put forward in this thesis. Both Percus (1997) and I argue that the cleft clause restrictively modifies the constituent *it*, which functions both as the definite article and a semantically underspecified head noun. However, for Hedberg, the constituent *it* functions as the determiner, with the cleft clause corresponding to the entire nominal content of the determiner phrase. The differences in these analyses seem to result from a difference in approach. While Hedberg (2000) examines *it*-clefts in relation to definite descriptions as they occur outside of the specificational construction, Percus (1997) and I compare *it*-clefts to other specificational copular sentences containing definite noun phrases.

As a result of our common approach, Percus (1997) and I recognize that *it*-clefts correspond closely to *th*-clefts, which are introduced by restrictively modified, semantically general head nouns such as *one* and *thing*. From this, it follows that the cleft *it* functions as both the definite article and head noun of a definite description and is modified by a restrictive relative clause. Hedberg (2000), on the other hand, does not make this connection. She suggests that the definite article is realized as *it* as a consequence of extraposition; that is, "when no descriptive content (NP or CP) immediately follows" (Hedberg 2000: 898). However, as I go on to show in §7.1, this is not supported by the historical evidence, since the pronoun *it* could sometimes occur with an immediately following restrictive relative clause at earlier periods of the language.

In Hedberg's analysis (2000: 912), the "complement clause" is extraposed from the subject DP and is subsequently adjoined to the focal element, within the VP, as shown in (75).

(74) It was Clinton who won (Hedberg 2000: 891)

(75) $[_{IP}\ [_{DP}\ [_{D}\ it]]_i\ [_{I'}\ [_{I}\ was_k]\ [_{VP}\ [_{DP}\ [_{DP}\ Clinton]\ [_{CPi}\ [who_j]\ [_{C'}\ t_j\ won]]]]]]$

Hedberg (2000) argues that this analysis represents the "best of both" the expletive and the extraposition accounts. By recognizing that the complement clause and the initial *it* function pragmatically and semantically like a discontinuous definite description, her analysis is able to capture the many similarities between *it*-clefts and definite noun phrases. However, on her account, the complement clause is syntactically adjoined to the clefted constituent. Consequently, this analysis also accommodates Delahunty's (1982, 1984) claim that the postcopular element and the cleft clause form a syntactic constituent within VP.

However, Hedberg's (2000) extraposition account is *ad hoc*. As I noted above, for Hedberg, the "complement clause" of the subject DP is extraposed into a position internal to another DP inside the verb phrase. The upshot is that the complement clause is adjoined to the postcopular DP "in the position of a nonrestrictive relative clause" (Hedberg 2000: 915). Unlike right-dislocation or extraposition-from-NP, this type of extraposition is novel to this construction and is therefore unmotivated. Furthermore, Hedberg's account is unnecessarily complex. As I go on to explain in §4.3, the data that Delahunty (1982, 1984) uses to argue in favour of VP-internal constituency is not problematic for a straightforward extraposition-from-NP analysis such as the one argued for in this thesis and does not necessarily require that the focal element and the cleft clause should be analysed as syntactic sisters.

Hedberg's (2000) syntactic analysis of *it*-clefts is therefore very different from that proposed here. However, semantically, there are some similarities. Like me, Hedberg does not provide an equative analysis for specificational *it*-clefts. Instead, she suggests that the postcopular element is referential (type <e>), while the definite

description is of a higher semantic type which is either a predicate (type <e,t>) or a generalized quantifier (type <<e,t> t>) (Hedberg 2000: 917). However, confusingly, elsewhere in her paper, Hedberg (2000: 891) claims that the constituent *it* in *it*-clefts is "referential", with the cleft *it* and the cleft clause forming a "discontinuous referring expression" (Hedberg 2000: 898). Furthermore, Hedberg's (2000: 916) semantic analysis relies on an "identificational copula"; this element is required to obtain a specificational (or identifying) rather than a predicational meaning and also plays an important role in assigning exhaustive identification to the clefted constituent. On my account, on the other hand, specificational meaning (as well as exhaustiveness) results from a nominal predication relation involving definite NP predicates and is not reliant on a special type of copular verb (see §3.3.3 and §4.1.2.3).

In her 1990 thesis, Hedberg also argues for a discontinuous constituent analysis of *it*-clefts. Although her argument is largely consistent with Hedberg (2000), Hedberg (1990) proposes a different syntactic structure for the *it*-cleft. For example, Hedberg (1990) claims that the cleft clause is structured internally like a restrictive relative clause. Likewise, rather than adjoining within DP, the cleft clause adjoins to VP, shown in (77). Nevertheless, like Hedberg (2000), Hedberg (1990: 99) assumes that the cleft clause forms a constituent with the postcopular element within VP, claiming that "the relation which holds between the cleft pronoun and the cleft clause is not identical to the relation which holds between an extraposed relative clause and its NP head". Again, this differs from my own straightforward, extraposition-from-NP account of *it*-clefts.[18]

(76) It is Joe that smokes

(77) $[_{IP}\ [_{NP}\ it]_j\ [_{I'}\ [_{I'}\ [_{I}]\ [_{V}\ is_k]]\ [_{VP}\ [_{VP}\ [t_k]\ [_{NP}\ Joe]]\ [_{CPj}\ [_{OPi}]\ [_{C'}\ [_{C}\ that]\ [_{IP}\ t_i$ smokes]]]]]]

[18] Reeve (2008) proposes a syntactic account of *it*-clefts which draws heavily from Hedberg (1990, 2000). He claims that the cleft clause is a restrictive relative which is underlyingly adjoined to the clefted constituent within DP. The cleft clause is subsequently extraposed and is adjoined to VP, forming a constituent within VP. Semantically, however, Reeve argues in favour of Percus' (1997) analysis, in which the restrictive relative clause and the non-expletive *it* function together like a discontinuous definite NP and the copula has an equative semantics.

4.2.9 Han and Hedberg (2008)

In subsequent, related work, Han and Hedberg (2008) propose a discontinuous constituent analysis for *it*-clefts within Tree Adjoining Grammar. Again, like Hedberg (2000), their aim is to "capture the best of both" extraposition and expletive accounts (Han and Hedberg 2008: 345). In order to do this, Han and Hedberg exploit the fact that Tree Adjoining Grammar allows for mismatch between the syntax and the semantics. As part of the machinery of Tree Adjoining Grammar, structures are analysed both with a derived tree, which represents surface constituency, and a derivation tree, which represents semantic composition and syntactic dependencies (Han and Hedberg 2000: 355). In the derived tree, the cleft clause (CP) is syntactically adjoined to the focal phrase (FP), forming a small clause (FP). However, in the derivation tree, the pronoun *it* and the cleft clause form a syntactic and a semantic unit, which is represented by placing elementary trees for them in a single multi-component set.

Han and Hedberg (2008) claim that this analysis explains the fact that *it*-clefts share a number of properties with definite descriptions, while at the same time accounting for the evidence in favour of an expletive account. However, Han and Hedberg's (2008) derived tree does not straightforwardly explain the connectivity effects (discussed in §4.3) which support an expletive analysis, since the postcopular element is not situated within the cleft clause at any stage in the derivation. In order to accommodate this awkward data, Han and Hedberg make use of feature unification. For example, they postulate an agreement feature attribute (Agr) in order to capture agreement between the cleft pronoun and the copular verb and between the focal element and the verb in the cleft clause respectively. Essentially, all that Han and Hedberg's (2008) analysis really buys us then, is a way of accommodating Delahunty's syntactic constituency data into an extraposition account. However, as I explain in §4.3, this data is not actually problematic for the straightforward extraposition-from-NP analysis proposed in this thesis.

As a result, I believe that Han and Hedberg's (2008) "two-tree" analysis for *it*-clefts is unnecessarily complex. Han and Hedberg claim that the advantage of a Tree Adjoining Grammar model is that is allows for mismatch between the semantics and the

syntax. Since my own constructional analysis involves mismatch, I have no theoretical objection to this. However, while in my analysis, mismatch is motivated and explained as a consequence of the nature of semantic predication, Han and Hedberg provide no real justification as to why, in their analysis, the syntax and semantics of *it*-clefts should be so completely at odds with one another. To my mind, it seems as though Han and Hedberg make opportunistic use of the "two tree" model in order to propose two disparate analyses of *it*-clefts within the same proposal.

Semantically, Han and Hedberg (2008) argue for an equative analysis of specificational *it*-clefts, in which an "equative copula" identifies two referential arguments: the discontinuous constituent and the postcopular element.

(78) It was Ohno who won

(79) THE*z* [won(z)] [z = Ohno] (Han and Hedberg 2008: 349)

Consequently, Han and Hedberg's (2008) analysis suffers from the same problems as other equative analyses, such as Heycock and Kroch (1999, 2002) (see §3.2). I discuss this equative analysis in more detail in §5.2.3, where I review Han and Hedberg's (2008) account of predicational *it*-clefts.

4.2.10 Summary and interim conclusions

In this section, I have shown that, while the analysis of *it*-clefts proposed in this thesis builds upon many of the insights of earlier extraposition accounts, it also improves upon the existing literature. None of the extraposition accounts I have reviewed here recognize the crucial role that definite descriptions play in creating specificational meaning. It is this that provides the most compelling support for a discontinuous definite description analysis of *it*-clefts. My analysis examines the semantics of definite NPs in more detail than other discontinuous constituent accounts, which do not recognize that definite noun phrases in specificational sentences, and the discontinuous constituent in *it*-clefts, denote sets. The set treatment of definite descriptions is the basis for an analysis in which specificational sentences involve a class-membership predication relation.

4.3 Remaining questions and problems

In this section, I address some of the difficult data which many have claimed is at odds with an extraposition analysis of *it*-clefts. For example, Delahunty's (1982, 1984) constituency tests and number agreement issues at first sight support an expletive analysis of *it*-clefts. However, as I explain below, this data does not present an insurmountable obstacle to an extraposition analysis and can be accounted for on this approach. I also discuss Jespersen's (1937) criticisms of his own earlier "transposition" analysis, noting that many of these do not apply to the extraposition account argued for here. Finally, I briefly address the binding (or "connectivity") effects which have been used to support analyses in which specificational copular sentences are derived via movement operations. I conclude that although the account proposed in this thesis is not without its problems and leaves some questions unanswered, it has many advantages over the alternatives.

4.3.1 Constituency tests

As I noted in §4.2.8, Delahunty (1982, 1984) argues that the postcopular focal element and the sentence-final clause of *it*-clefts form a constituent within VP. This claim is at odds with the straightforward extraposition-from-NP analysis advocated here. Delahunty (1982) presents five constituency tests in support of his analysis, including VP-deletion, right-node-raising, parenthetical formation, VP-conjunction and VP-fronting. With respect to each of these environments, shown in (80) to (84), Delahunty claims that the focal element and the cleft clause operate as a syntactic unit (underlined).

(80) *VP deletion:* I said that it should have been Bill who negotiated the new contract, and it should have been.

(81) *Right-node-raising:* It could have been – and it should have been – Bill who negotiated the new contract.

(82) *Parenthetical formation:* It must have been, in my opinion, the cyanide that did it.

(83) *VP-conjunction:* It must have been Fred that kissed Mary but Bill that left with her.

(84) *VP-preposing:* I said that it was Bill that argued the case, and Bill that argued the case it was. (examples adapted from Hedberg 1990: 98)

There are three possible responses to Delahunty's claims. First, this data is not necessarily problematic for my account. Second, Delahunty's constituency tests are not necessarily reliable. And third, there is constituency evidence for a discontinuous definite description analysis. I take these in turn, starting off by discussing how these data do not all constitute arguments against my analysis.

I argued in §4.1 that the element *it* in *it*-clefts is restrictively modified, providing the same function as the definite article and head noun of a definite description. However, in so-called "truncated clefts", the initial *it* is a full pronoun which is anaphoric to a complete set description. From this it follows that example (85) represents a separate specificational construction from the *it*-cleft and does not involve the deletion of the restrictive relative clause.

(85) A. Is John **the murderer**? B. No, **it**'s Gary.

In examples of VP-deletion then, such as (80) (repeated here as (86)), it is consistent with the analysis proposed in this thesis that the second instance of *it* is a full pronoun (presented in bold) and is anaphoric to the description *it (the one) who negotiated the new contract* (underlined). On this analysis, the example in (86) involves the "deletion" only of *Bill* and not *Bill who negotiated the new contract*. Consequently, this data is not problematic to an extraposition analysis which recognizes that, in specificational sentences, *it* can operate both as a full pronominal NP and as the head of a restrictively modified definite description. As a result, VP-deletion cannot be used as an argument against the analysis proposed in this thesis.

(86) I said that it should have been Bill who negotiated the new contract, and **it** should have been.

Now we can turn to the issue of why these constituency tests are not all reliable. As Croft (2001: 189) comments, coordination is not a reliable test for constituency. This calls into question the "constituency" results of right-node-raising, shown in (81) and VP-conjunction, shown in (83). For example, Croft (2001: 189) provides the following examples of English coordinate constructions which "support "constituents" that are not supported by other criteria".

(87) *Right-node-raising:* [Jenny makes]$_{??}$ and [Randy sells] $_{??}$ the prints

(88) *Gapping:* Jenny gave [the books to Randy]$_{??}$ and [the magazines to Bill]$_{??}$

(Croft 2001: 189)

In line with Croft's (2001: 189) observations about the unreliability of coordination-based constituency tests, Hedberg (1990: 98) notes that right-node-raising, parenthetical formation and VP-conjunction constructions "do not entirely exclude extraposed relative clauses…from appearing in the position filled by the cleft clause in clefts". She provides the following examples. In each of these cases, the extraposed relative clause clearly restricts, and forms a "discontinuous constituent" with, *nobody*, but shows the same pattern of behaviour as the *it*-clefts given in (81) to (83) above.

(89) Nobody would – and nobody could – drink instant coffee who knew anything about espresso.

(90) Nobody would ever, in my opinion, drink instant coffee, who knew anything about espresso.

(91) Nobody could drink instant coffee and enjoy it, who knew anything about espresso. (Hedberg 1990: 99)

This evidence therefore shows that the data involving right-node-raising, parenthetical formation and VP-conjunction is equally consistent with my extraposition-from-NP account of *it*-clefts and does not undermine my analysis.

This leaves us with VP-fronting (or VP-preposing). Hedberg (1990: 98-99) shows that VP-preposing is not possible for extraposed relative structures, giving the example shown here as (92). On this basis, Hedberg claims that *it*-clefts do not involve straightforward extraposition-from-NP, and the cleft clause must instead be a constituent of the VP.

(92) *I said a candidate would win who had charisma, and win who had charisma, a candidate did. (Hedberg 1990: 99)

However, even Delahunty (1982, 1984) observes that VP-fronting in *it*-clefts, shown in (84), is unacceptable to many speakers. In his (1982) thesis, he builds up his analysis that the postcopular element and the cleft clause are syntactic sisters without this evidence, treating such examples as ungrammatical. Likewise, Delin (1989) considers the ungrammaticality of preposing in (93) to be evidence against treating the postcopular element and the cleft clause as a constituent.

(93) *I said it is John that's an interesting guy, and John that's an interesting guy it is. (Delin 1989; cited in Hedberg 2000: 916f)

Nevertheless, in his (1984) paper, Delahunty argues that the low acceptability of *it*-cleft examples involving VP-fronting is dependent upon "stylistic awkwardness" rather than ungrammaticality. Similarly, Hedberg (1990: 98) claims that since some of these examples are more acceptable than others, preposing can be used as evidence for constituency. She notes that the *it*-cleft given here as (94) and the noncleft example of VP-preposing in (95) are "equally acceptable".

(94) I said that it would be a conservative who'd win, and a conservative who won it certainly was.

(95) I said that I would finish by September, and finish by September I did.

(Hedberg 1990: 98; her judgments)

However, I disagree with Hedberg (1990). The *it*-cleft in (94) is clearly less acceptable and is certainly less natural than the noncleft example given in (95).

In conclusion then, examples of VP-preposing are at best unnatural and at worst ungrammatical. Consequently, I do not view this data as strong evidence for VP-internal constituency. In sum, Delahunty's constituency tests are not problematic for an extraposition analysis of *it*-clefts, since this data (where grammatical) can be accommodated fairly neatly.

We can follow Delahunty's discussion by pointing out that there is evidence that the cleft clause does **not** form a constituent with the postcopular element. For example although Han and Hedberg (2008:357) conclude that the clefted constituent and the cleft clause form a syntactic constituent, they note that they can be separated by an adverbial phrase, shown in (96).

(96) It was Kim, in my opinion, who won the race (Han and Hedberg 2008: 357)

This adverbial phrase records speaker opinion rather than the manner of the act denoted by the verb. As a result, *in my opinion* must be a sentence adverbial rather than a VP adverbial and cannot be located within the VP. The fact that a sentence adverbial can come between the clefted constituent (*Kim*) and the cleft clause (*who won the race*) provides a solid argument against the claim that they are syntactic sisters within VP.

4.3.2 Agreement issues

Moving beyond constituency diagnostics, we can look at agreement facts, which have also been used to argue against an extraposition analysis of *it*-clefts (see for example Jespersen 1937). I argue here that, for the most part, *it*-cleft agreement actually supports

an approach in which *it*-clefts are analysed as a subtype of the specificational copular construction. For example, Akmajian (1970) notes that, for the vast majority of speakers, the verb in the cleft clause is systematically marked as third person and pronouns in postcopular position are always in the objective case, shown in (97).

(97) It's **me** who **does** this job (Akmajian 1970: 151)

This pattern is to be expected on an extraposition account, since the *it*-cleft in (97) corresponds to the pseudocleft in (98). Here, the verb in the relative clause shows agreement with the third person head noun *one* and the focal element is marked as objective as a result of its postverbal position.

(98) The **one** who **does** this job is me

However, this data is problematic for expletive accounts which assume a predication relationship between the clefted constituent and the cleft clause. Under an expletive analysis, we would expect the postcopular element in (97) to be marked as nominative (since it acts as the preposed argument of the proposition expressed in the cleft clause) and for the verb in the cleft clause to show person agreement with the clefted constituent, as shown in (99). However, this pattern is not found in the dialect of the majority of speakers.

(99) *??It is **I** who **do** this job

The person agreement data is particularly awkward for expletive analyses which invoke movement to extract the clefted constituent out of the cleft clause (including Rochemont 1986), since examples such as (97) do not have corresponding noncopular paraphrases. In order to accommodate this data, É. Kiss (1998: 259), whose main argument requires movement, suggests that the clefted constituent can also be base-generated in the postcopular position and coindexed with the *wh*-pronoun in the cleft clause.

Nevertheless, although expletive analyses invoking the base-generation strategy can accommodate this person agreement data, they still cannot explain it.

While person agreement and case provide evidence in support of an extraposition analysis, number agreement is much more problematic. As shown in (100), the verb in the cleft clause agrees in number with the clefted constituent. Akmajian (1970) claims that the number agreement patterns in *it*-clefts are also found in their corresponding pseudocleft sentences. For example in (101) the plural verb (*are*) in the relative clause agrees in number with the head noun of the definite description (*ones*).

(100) It is John and Sally [who **are** responsible]

(101) The ones [who **are** responsible] are John and Sally

However, as Huddleston (1984: 461) comments, there are some important differences in the number agreement patterns of *it*-clefts and pseudoclefts. For instance, from (100) above we can see that the copular verb in the matrix clause of the *it*-cleft is always singular and does not agree in number with the verb of the relative clause. In contrast, in the pseudocleft in (101), both the copular of the matrix clause and the verb in the relative clause agree in number with the head noun *ones*. This observation is problematic for my analysis of *it*-clefts. From the example given in (100) it would appear that while the copular verb of the matrix clause agrees in number with the singular pronoun *it*, the verb of the relative clause shows agreement with the focal element, exactly as would be predicted on an expletive account.

Nevertheless, the semantic account of specificational sentences proposed in this thesis can help to explain this data. I argued in §3.3 that specificational copular sentences involve a nominal predication relation in which we are told the membership of a restricted set. In §4.1, I extended this analysis to *it*-clefts, arguing that the discontinuous definite description, made up of *it* and the restrictive relative clause, denotes a restricted set. In pseudoclefts such as (102), the restricted set is described by its (plural) members, *the **ones** that travel by train*. The *it*-cleft construction can also identify more than one member as comprising a restricted set. However, in this case,

shown in (103), the constituent *it* and the copular verb are marked as singular. Effectively, in this example, a singular set (*of co-workers that travel by train*) is identified as having plural members.

(102) The ones that travel by train **are** John, Suzie and Tom

(103) It**'s** John, Suzie and Tom that travel by train

In this way, *it*-clefts behave differently from specificational sentences with overt nominal heads. It is possible that this small difference could well lead to the interesting number agreement patterns we find in *it*-clefts. In construction grammar, it is accepted that speakers may have different conceptualizations underlying a single linguistic structure and that this can lead to variation and change. For example, Croft (2001: 127) provides the following attested example in which a noun marked as singular is followed by a plural verb. Croft explains this phenomenon, which has become a convention of the language, as the result of speakers having alternative conceptualizations available underlying collective entities, such as *a highway authority*. He notes that "a collective entity is conceptually both singular and plural; it is singular because it functions as a singular unit, but it is also plural in that it is made up of a multiplicity of individuals" (Croft 2001: 128).

(104) Section 278 of the Highways Act 1980, therefore, provides that if **a** highway authority **are** satisfied that it would be of benefit to the public…

(Croft 2001: 127)

Since, in the analysis presented in this thesis, *it* and the relative clause denote a set, there are also two possible conceptualizations of this entity. In (103), this discontinuous description is both singular (in relation to the set) and plural (in relation to its members). Consequently, it is possible that while the copular verb of the matrix clause shows agreement with the singular set, marked as singular by *it*, the verb in the relative clause shows number agreement with the plural members of this set. This might

sound opportunistic, but it is worth noting that a similar state of affairs is shown in *NP of NP* constructions. In the specificational copular sentence in (105), the singular noun *group* agrees with the matrix verb while the plural noun *students* agrees with the verb in the relative clause.

(105) The group of **students** that **are** always late is John, Suzie and Tom

If this is the correct analysis for number agreement in *it*-clefts, we would expect that the extraposition of the relative clause plays a fundamental role in the lack of agreement between the initial *it* and the verb in the relative clause. For example, in (106), the restrictive relative clause is given after the identification of the members of the set. This tells us that the set denoted by the discontinuous constituent is a singular set with plural members and allows the plural marking of the verb in the relative clause.[19] If this is the case, then in (106), we are saying that *the restricted set (of students/individuals that are always late) is comprised of John, Suzie and Tom.*

(106) It is John, Suzie and Tom that are always late

Tellingly, there is also variation in this pattern. For some speakers, the verb in the relative clause can be marked as singular (agreeing with *it* and the copular verb) even when the clefted constituent is plural. For instance, in my dialect (107) is an equally acceptable alternative to (106) above.

(107) It is John, Suzie and Tom that's always late

[19] Like *it*-clefts, *wh*-clefts can also occur with plural foci. As with the *it* of *it*-clefts, *what* is inherently singular and shows agreement with the copular verb of the matrix clause. However, for *wh*-clefts, the verb in the relative clause is also marked as singular regardless of the plurality of members, shown in (i). Again, this suggests that it is the extraposition of the relative clause in *it*-clefts that leads to the lack of agreement between the initial *it* and the embedded verb.

(i) **What is** on the table **is** an egg and a screwdriver

In conclusion then, number agreement is a problem for an extraposition analysis of *it*-clefts. Nevertheless, as we have seen, agreement patterns on the whole (including person and case marking) seem to favour this analysis over an expletive account. Furthermore, the number agreement patterns gain at least a plausible explanation in the particular extraposition account proposed in this thesis, since the initial *it* and the relative clause together form a singular set (with an underspecified head noun). From what we know about the different possible conceptualizations of collective entities, it is possible that while the copula of the matrix clause shows agreement with the singular set, the relative clause provides restrictions on the members of this set, resulting in plural marking for this verb. In §5.2.2, I show that this analysis is able to account for the different agreement patterns of specificational and predicational *it*-clefts. I provide historical evidence in support of my account of the *it*-cleft's number agreement patterns in §7.1.3.

4.3.3 Jespersen's (1937) critique

I have shown then that Delahunty's (1982, 1984) constituency tests and the number agreement puzzles outlined above can be accounted for within my extraposition-from-NP analysis of *it*-clefts. However, Jespersen (1937) raises a number of further objections to an extraposition account of *it*-clefts, which were originally proposed as a means of criticizing his earlier transposition theory (see §4.2.1). I discuss some of these points here and explain that many of them do not represent a serious problem for my analysis. Aside from the number agreement issues discussed above, Jespersen's (1937) criticisms lie mainly in presenting data pertaining to the nature of the cleft clause. For example, Jespersen (1937: 84) notes that the cleft clause prefers *that* to *who* or *which*, can occur without a relative marker, and is not separated from the rest of the sentence by commas or pauses. On this basis, Jespersen argues that the cleft clause and postcopular element form a constituent. However, each of these points is discussed in §4.1.3 where I show that they support the analysis of the relative clause in *it*-clefts as restrictive rather than nonrestrictive. Consequently, they do not pose a problem for the account of *it*-clefts proposed in this chapter.

In fact, Jespersen (1937: 84-85) seems to be arguing against a different version of the extraposition approach from that which is proposed in this thesis. For example, he notes that the transposition theory cannot "easily account for those cases in which it is…an adverb or a similar word that follows after *it is*", since (108) is not equivalent in meaning to (109). However, my analysis involves extraposition-from-NP rather than extraposition of the NP. Consequently, the *it*-cleft in (108) is not analysed as a variant of (109), but as corresponding in meaning to the specificational copular sentence in (110). As a result, Jespersen's criticism does not apply to the analysis proposed in this chapter but to accounts in which the entire subject NP is extraposed.

(108) It was here that he died (Jespersen 1937: 85)

(109) That he died was here (Jespersen 1937: 85)

(110) The place that he died was here

However, Jespersen (1937: 84-85) does raise two important points which represent a problem to the analysis I have built up in this chapter. First, he notes that the cleft clause and the clefted constituent are always adjacent to one another. This raises the question: if the relative clause modifies the constituent *it*, then why can't it occur immediately after the initial *it* in *it*-clefts? Second, he notes that unlike the ordinary antecedents of relative clauses, *it* cannot be stressed. I deal with these two issues in §7.1 where I show that the present-day *it*-cleft construction has retained an Old English paratactic relative clause structure. Unlike other restrictive relative clauses, the cleft clause has never come together with its antecedent, possibly as a result of information-structure tendencies and/or prosodic factors.

4.3.4 Connectivity effects

The final set of difficult data involves connectivity effects. This data has been used to argue in favour of a movement-based, derivational account of cleft sentences. While it is beyond the scope of this thesis to provide a constructional theory of binding, I hope to

show that the connectivity data actually presents a problem, rather than evidence, for an analysis involving movement.

In generative and transformational theories of grammar, binding effects represent clause-internal structural relations between syntactic elements. On this account, it follows that the clefted reflexive *himself* in (111) is somehow c-commanded by the pronoun *he* embedded inside the cleft clause. On the face of it, binding in *it*-clefts seems to support an expletive account, in which the clefted constituent is either moved out of the cleft clause or is coindexed with a gap inside the cleft clause. Such an analysis allows binding in *it*-clefts to work in much the same way as it does in simple noncopular sentences, such as (112) (see Hankamer 1974).

(111) It was **himself** that **he** hurt

(112) **He** hurt **himself**

However, Higgins (1979) raises an important problem with this analysis. He notes that the same connectivity effects are found not only in *it*-clefts, but in all types of specificational copular sentence. For example, in the *wh*-cleft and the noncleft specificational sentence below, the reflexive *yourself* in the matrix clause is governed by the pronoun *you* embedded inside the restrictive relative clause.

(113) What **you** should try instead is shaving **yourself** in the evening

(114) The approach **you** should try instead is shaving **yourself** in the evening

(examples from Higgins 1979: 56)

Higgins therefore calls for a unified account of specificational copular constructions. Since noncleft examples such as (114) cannot be derived from simple noncopular sentences (such as ***you*** *should try shaving* ***yourself*** *in the evening*), Higgins claims that

the use of reflexive pronouns in specificational sentences requires an independent explanation from syntactic c-command.[20]

In specificational copular sentences then, reflexives in the focal XP are not c-commanded by the elements on which they depend for their interpretation; that is, the two pronouns are not located within the same clause. The connectedness in specificational sentences therefore raises interesting questions about how and where binding is licensed.

Of course, specificational sentences are not the only examples which challenge the claim that binding relies on structural c-command. For example, in the noncopular sentence in (115), *John* does not c-command the reflexive *himself*. Such sentences involve "backward binding" into the subject NP. For Williams (1983, 1994), a proponent of the inverse analysis of copular constructions, binding in specificational sentences works in much the same way. On the inverse account, specificational sentences, such as the *wh*-cleft in (116), are derived via movement from subject-predicate structures such as (117). Williams therefore claims that although binding in (116) superficially runs from left to right, it is nevertheless derived from a structure involving backward binding into the (now postcopular) subject.

(115) A picture of himself$_i$ upset John$_i$ (Heycock and Kroch 1999: 369)

(116) What he$_i$ promised was to reform himself$_i$ (Heycock and Kroch 1999: 368)

(117) To reform himself$_i$ was what he$_i$ promised

However, Heycock and Kroch (1999: 369) argue that backward binding does not present a viable explanation for the binding effects of specificational copular sentences. They note that while the connectivity effects in specificational sentences are categorical, backwards binding appears to be optional (see Heycock and Kroch 1999: 369-370).

It seems then that connectivity in specificational copular sentences cannot be satisfactorily accounted for by movement-based analyses. Alternative explanations for

[20] Higgins (1979) frames this discussion in response to Akmajian (1970), who claims that while *it*-clefts are derived from their pseudocleft sources via extraposition, these are in turn derived from simple noncopular sentences which conform to standard reflexivization rules.

these connectivity effects include the question-in-disguise theory of pseudoclefts championed by Den Dikken *et al.* (2000) and Schlenker (2003), the semantic accounts of Jacobson (1994), Sharvit (1999, 2003) and Heller (2002, 2005) and the information-structure based accounts of Heycock and Kroch (1999, 2002) and Lahousse (2009).

It is beyond the scope of this thesis to review these alternative theories in detail here. However, as I see it, a problem with the "concealed question" analysis is that it stipulates enormous amounts of elided material which does not seem to be independently motivated or supported by agreement factors.[21] Likewise, most semantic accounts, which use higher-order semantics rather than syntactic c-command to license connectivity, are based upon the premise that specificational sentences are identity statements (see for example, Sharvit (1999, 2003) and Heller (2002, 2005)).[22] I have already presented a number of reasons for rejecting an equative analysis of specificational sentences in §3.2. Finally, the information-structure based account proposed by Heycock and Kroch (2002) struggles to account for the difficult data for which it was intended. These authors claim that canonical order specificational sentences share the same 'focus/ground' information structure as noncopular simple sentences, which in turn accounts for their shared connectivity effects. However, the explanatory power of this analysis is dependent upon the claim that reverse specificational sentences have a different information structure involving topicalization. This is not supported by the analysis of specificational sentences argued for in this thesis.[23]

---

[21] An important advantage of the question-in-disguise or "ellipsis" approach is that it can account for the licensing of negative polarity items in examples such as (i) by examining them in relation to *wh*-clefts with sentential foci, such as (ii), in which the NPI is licensed by a c-commanding pronoun. However, as I explain in §8.2, examples such as (ii) are instances of a separate 'presentational' *wh*-cleft construction. Consequently, this structure should not be used as the basis for an analysis of *wh*-clefts in general.

(i) What **he** didn't buy was **any** wine
(ii) What he didn't buy was, **he** didn't buy **any** wine

[22] Jacobson (1994) frames her semantic account of connectivity within the inverse analyses of Williams (1983) and Partee (1986). In theory then, this analysis should be broadly compatible with my own predicational account of specificational sentences. However, Heycock and Kroch (2002: 380f) argue that "Her approach seems instead to require the equative analysis that we propose".

[23] Heycock and Kroch (2002) attempt to account for the differing connectivity properties of canonical and reverse specificational sentences (shown below) by suggesting that while reverse specificational sentences

I have shown that the exact nature of these connectivity effects is elusive and I do not attempt to propose a constructional theory of binding here. However, it is my belief that the analysis of specificational meaning provided in this thesis is not at odds with the connectivity data and could actually help to make sense of it. In what follows, I sketch a very simple account of reflexivization in *it*-clefts and other specificational *NP be NP* sentences which, crucially, explains why these connectivity effects are not found in predicational *NP be NP* sentences.

We can start with the very simple premise that the participant roles of a predicate must be reflexively marked if they correspond to the same entity. In the *it*-cleft in (118), the embedded verb *hurt* takes two participant roles, an agent and a patient. The purpose of this specificational sentence is to identify the patient, described by the discontinuous definite NP (*the one that he hurt*). In other words, the postcopular element provides the referent which matches the NP-internal participant of the embedded verb *hurt*. Since the patient and the agent roles are performed by the same individual, a reflexive in the postcopular position is expected.

(118) It was **himself** that **he** hurt

This same kind of semantic analysis can be extended to account for other specificational copular sentences. For example, the specificational *wh*-cleft in (119) identifies a property of *Sarah* rather than a 'participant'. Nevertheless, this example can be accounted for by the same set of assumptions i.e. that nouns can only assign theta-roles within the NP (see Francis 1999a, b) and that connectivity effects result ultimately from the semantics of specification. In (119), the noun *burden* assigns participant roles

---

can have a focus/ground information structure, the precopular NP is often discourse old, resulting in a topic/focus information structure. There are two problems with this analysis. First, their account predicts that if (ii) is used to answer the question *what did he miss the most?*, resulting in primary stress on the head noun *dog*, then the co-reference reading should not be possible. This, however, is not the case. Second, I have argued in §3.3 that reverse specificational sentences **always** have a focal subject. Otherwise, the result is a non-specificational instance of the predicate nominal construction.

(i) What $he_{*i/j}$ really missed was $John_i$'s dog
(ii) $John_i$'s dog was what $he_{i/j}$ really missed (examples from Heycock and Kroch 2002: 119)

within the postcopular NP. It is a "participant nominalization", meaning that the noun *burden* incorporates the participant role of agent as well as assigning the patient role to *herself* (see Francis 1999b).

(119) What **Sarah** is is a burden to **herself** [specificational *wh*-cleft]

On this account then, binding is internal to the NP. The only remaining question is how do we identify the agent participant, (*the one who is*) *a burden*, with *Sarah*? This comes about via the identification of the postcopular referent with the head noun of the precopular description (*what*). Since *what* is marked as denoting a property by the restrictive relative *Sarah is x*, it follows that in (119), *Sarah* is classified as *a burden*.

This relatively simple, semantic account of reflexivization explains why the same connectivity effects are not found in predicational copular sentences, such as (120). On this predicational reading, *what Sarah is*, refers to a property, such as *being a perfectionist*, which is described or classified as *a burden to her*. Here, the proper noun *Sarah* and the non-reflexive pronoun *her* co-refer. This suggests that these entities are not both participants of the same predicate nominal.

(120) What **Sarah** is is a burden to **her**/*herself [predicational *wh*-cleft]

As in (119) above, *burden* is a participant nominalization which incorporates the role of agent. By classifying the referent *what Sarah is* as *a burden to her*, both the entity referred to in the precopular NP and the agent participant in the postcopular NP are identified as the same individual. However, the agent and patient participants of the postcopular NP (*what Sarah is* and *Sarah*) do not correspond to the same entity (shown in (121)). As a result, the second participant cannot be marked as reflexive.

(121) That$_i$ is a burden to her$_j$/*herself$_i$ [predicational *NP be NP*]

On this semantic account of reflexivization, binding in both specificational and predicational copular sentences is internal to the NP. While the class-membership predication relation may give the appearance of syntactic connectivity across clauses, this is simply caused by matching a referent to the NP-internal participant of a predicative element. These predicative NPs correspond to simple noncopular sentences. For example, the participant nominalizations *a burden to herself* and *a burden to her* in (119) and (120) above correspond to *she bothers herself* and *that bothers her*, respectively. This analysis therefore explains why there is a perceived difference in connectivity between specificational and predicational copular sentences (see Sharvit 1999), but also manages to provide them with a unified analysis.

4.3.5 Summary and interim conclusions

Throughout §4.3, I have shown that while the analysis of *it*-clefts and specificational copular sentences proposed in this thesis may leave some questions unresolved, much of the difficult data can be accommodated into this account. On the whole, I believe that my analysis has many important advantages over alternative accounts and not too many disadvantages. For example, one advantage is that my analysis of *it*-clefts has explanatory power, maximizing motivation from the language system. In the next chapter, I highlight the areas that are beyond the limits of what can be explained as a consequence of inheritance from the wider specificational copular construction. I ask whether these seemingly idiosyncratic properties can be shown to be motivated in some other way, thereby extending the scope of my analysis.

5. EXTENDING THE CONSTRUCTIONAL ACCOUNT

Over the previous two chapters, I have shown that *it*-clefts can be integrated into an analysis of specificational copular sentences. In chapter 3, I argued that specificational meaning arises from a nominal predication relation involving definite noun phrase predicates. In chapter 4, I brought *it*-clefts into this account, arguing for a "discontinuous constituent" analysis, in which the initial *it* and the cleft clause work together to make a definite description. I showed that many of the *it*-cleft's properties are inherited from the wider specificational construction. For example, the pragmatic properties of focus, existentiality, exhaustiveness, contrast and givenness were explained either as the product of specificational meaning or as deriving from the behaviour of definite noun phrases. Likewise, the cleft clause was shown to be neither a non-modifying predicate (behaving like a relative clause as a result of analogy) nor a structurally unique form of relative clause (exhibiting both nonrestrictive and restrictive characteristics). Instead, by assuming that the cleft clause modifies the constituent *it*, its behaviour was shown to be consistent with a restrictive relative clause analysis.

Although this analysis has therefore been shown to have explanatory adequacy, it is nonetheless incomplete. By focusing on the properties that the *it*-cleft shares with the other members of the family of specificational copular constructions, the discussion so far has centred on what I will refer to here as the 'prototypical' *it*-cleft subtype, that is *it*-clefts that have a specificational meaning, contain nominal foci, and involve given information in the relative clause. However, in chapter 1 of this thesis, I noted that there are in fact several different subtypes of *it*-cleft, including predicational and proverbial *it*-clefts, non-NP focus *it*-clefts and 'informative-presupposition' *it*-clefts. For now, I assume that such examples represent non-prototypical subtypes of an overarching *it*-cleft schema. I explore this idea in chapters 7 and 8 of this thesis, where I use historical evidence to explain this family resemblance story.

In this chapter, I examine how these non-prototypical *it*-cleft subtypes are treated in the cleft literature and show that many analyses cannot accommodate them. I claim that an important advantage to my own analysis, outlined in §4.1, is that it can

incorporate all of the different types of *it*-cleft. However, although my analysis can accommodate the full range of *it*-cleft data, it cannot, as yet, explain it. Throughout this chapter, I show that predicational *it*-clefts, non-NP focus *it*-clefts and informative-presupposition *it*-clefts all exhibit properties that cannot be explained as a result of inheritance from the wider specificational copular construction. I ask whether we can find motivation for these properties that goes beyond inheritance from the specificational schema. I claim that while we can provide motivation for predicational and proverbial *it*-clefts owing to inheritance from the predicate nominal construction discussed in §3.3.3, we must look outside of the present-day language system to explain the range of *it*-cleft foci and the information-status of the cleft clause.

In §5.2, my discussion focuses on predicational and proverbial clefts. Here I show that expletive analyses of *it*-clefts cannot handle this data, I show that these examples can be integrated neatly into my account and I show how my account compares favourably to other attempts at accommodating these examples into an *it*-cleft analysis. In §5.3 I turn to non-NP focus *it*-clefts and in §5.4 I discuss the informative-presupposition subtype. In the next section, §5.1, I provide an introduction to each of these non-prototypical types of *it*-cleft and present a fuller description of the contents of this chapter.

5.1 The different types of *it*-cleft

In chapter 4, I focused on *it*-clefts that share the maximum number of properties with specificational *NP be NP* sentences containing definite noun phrase predicates. For example, the *it*-cleft in (1) is specificational, identifying *the therapist* as *the person that killed her*. Likewise, it contains a focal NP *the therapist*, which is used here as a referring expression. Finally, the relative clause expresses discourse-old information; that is, we know from the context that a woman has been murdered by someone. As I explained in §4.1.2.5, this property can be traced back to the semantics of definite noun phrases: since they involve existential commitment, whereby the entity described is assumed to exist, definite NPs are associated with given or shared information. Consequently, all three of these properties (specificational meaning, NP foci and given

information in the relative clause) are shared with specificational *NP be NP* sentences containing definite descriptions.

(1) A: Is he the murderer?
B: No. It was the therapist that killed her

However, there are other types of *it*-cleft which do not necessarily exhibit all of these three properties. These include predicational and proverbial *it*-clefts, non-NP focus *it*-clefts and informative-presupposition *it*-clefts.

5.1.1 Non-specificational it-*clefts*

There are *it*-clefts that have a predicational rather than a specificational meaning. These are known as predicational *it*-clefts and proverbial *it*-clefts. In these examples, the postcopular element provides descriptive rather than identifying information. For instance, (2) describes *the man that left* as *a tall man*. Likewise, the proverbial example in (3) tells us that *a road that has no turning is a long road.*

(2) It was a tall man that left [predicational *it*-cleft]
(3) It is a long road that has no turning [proverbial *it*-cleft]

As I explain in §5.2.1, such sentences have often been regarded as forming a structurally distinct sentence-type, which is only superficially related to the specificational *it*-cleft. However, I identify a number of problems with this claim and argue that these sentences should be integrated into a unified analysis of *it*-clefts. In §5.2.2, I show that my own particular analysis of *it*-clefts and of the specificational-predicational distinction is able to accommodate the predicational examples as well as providing an explanation as to why *it*-clefts permit both specificational and predicational readings.

In §5.2.3, I compare my analysis to that of Han and Hedberg (2008), who also extend their account to cover predicational (including proverbial) clefts. Like me, these

authors propose a "discontinuous constituent" analysis of the *it*-cleft. However, they suggest that the definite-like description and the postcopular element are both referring expressions. In this section, I present further objections to the claim that specificational sentences are semantically equative and show that my own "predicational" account of specificational constructions is better at explaining the different behaviour of specificational and predicational *it*-clefts.

5.1.2 It-*clefts with non-NP foci*

Another domain of variation concerns the range of elements that can occur as the complement of *be*. While noun phrases are the most common category to occur in the focal position, other possible categories include prepositional phrases and adverb phrases, among others.

(4) It's in December that she's coming [PP-focus *it*-cleft]

(5) It was just here that we met [AdvP-focus *it*-cleft]

These non-NP *it*-clefts are discussed in §5.3. Here, I show that the same range of potential foci is not shared by other kinds of specificational copular construction and therefore cannot be explained by inheritance. As a result, I claim that the *it*-cleft foci make up a construction-specific category and require an independent explanation. Since the factors that sanction and restrict the range of *it*-cleft foci do not apply wholesale to major syntactic categories, it follows that any attempt to define this construction-specific category in terms of general syntactic rules is unlikely to succeed. Instead, I argue in favour of a semantic description of *it*-cleft foci and suggest that historical evidence may help us to explain both what the boundaries of this construction-specific category are and how it has developed over time.

5.1.3 It-*clefts with new information in the cleft clause*

The final domain of variation of interest to us here concerns the discourse status of the cleft clause. As I noted above and in §4.1.2.5, the relative clause in *it*-clefts typically

expresses given, or hearer-old, information. However, some *it*-cleft examples contain hearer-new information in the relative clause. For instance, the *it*-cleft in (6) is used discourse initially. In this example then, the information in the relative clause is not given by the previous discourse and the proposition, that *someone once said 'laws are silent at times of war'*, is not even assumed to be shared knowledge.

(6) (Start of lecture)
It was Cicero who once said, 'Laws are silent at times of war'

I discuss these examples, which are commonly referred to as informative-presupposition *it*-clefts, in §5.4. Here, I show that the ability of the cleft clause to express hearer-new information cannot be inherited from the wider specificational copular construction, or in fact, from definite noun phrases – a claim which contrasts with the arguments of Gundel, Hedberg and Zacharski (2001). Instead, I suggest that these examples form a sub-construction that has emerged via extension from the prototypical *it*-cleft. Consequently, I argue that these examples require a diachronic explanation.

5.2 Predicational *it*-clefts

Predicational (and proverbial) examples present a particularly tricky analytical problem. Because they are less common than their specificational counterparts, predicational (and proverbial) *it*-clefts are rarely discussed in the cleft literature. As a result, for many authors, *it*-clefts are inherently specificational or identifying. For example, Gundel (1977: 547) suggests that the specificational-predicational distinction is not applicable to *it*-clefts, which "can only have an ID [identifying] interpretation". The question therefore arises, how do these predicational examples fit in with specificational *it*-clefts?

There are two possible solutions to this problem. One option is to continue to maintain that *it*-clefts are inherently specificational and to analyse the predicational examples as forming a distinct and unrelated sentence type. The other option is to acknowledge that *it*-clefts can have either a specificational or predicational meaning and to provide a unified analysis of both types of *it*-cleft.

I adopt the second of these strategies. In §5.2.2, I claim that, just as *NP be NP* sentences allow both specificational and predicational interpretations, so too do *it*-clefts. I outline an account in which predicational *it*-clefts inherit from the predicate nominal construction rather than from the specificational copular construction. I suggest that specificational and predicational *it*-clefts together form an overarching *it*-cleft schema, with the proverbial examples forming a subtype of the predicational *it*-cleft. I then go on, in §5.2.3, to explain how my analysis of the specificational-predicational distinction in *it*-clefts is more intuitive and economical than that proposed by Han and Hedberg (2008).

However, before I present my unified analysis of *it*-clefts, I review the accounts of those authors that choose the alternative strategy of providing non-specificational examples with a separate analysis from that of specificational *it*-clefts. In §5.2.1, I explain that since the expletive analysis of *it*-clefts cannot be extended to accommodate predicational examples, scholars who adopt this approach are forced to treat some or all predicational examples as forming an unrelated sentence-type. After outlining a number of problems with this reasoning, I conclude that the existence of the predicational subtype presents an important problem for expletive analyses of *it*-clefts.[1]

5.2.1 The expletive approach to predicational clefts

While Ball (1977), Declerck (1983) and Lambrecht (2001) all discuss non-specificational examples, they nevertheless maintain the claim that *it*-clefts are inherently specificational. For Ball (1977), predicational (including proverbial) clefts are "syntactically unrelated" to specificational *it*-clefts. Likewise, although Lambrecht (2001) discusses only the proverbial examples, he claims that they cannot be analysed as true "cleft" sentences. Declerck (1983) provides a somewhat different analysis, in that he distinguishes between proverbial and non-proverbial predicational clefts. Declerck (1983) argues that although proverbial clefts are superficially similar to the *it*-cleft proper, they actually form a structurally distinct sentence type. As for non-proverbial predicational clefts, Declerck (1983) claims that these can be integrated into his *it*-cleft

[1] The material in §5.2 was originally presented at the Cleft08 workshop in Berlin (see Patten 2008).

analysis; he treats them as "borderline cases", in which predicational information is contained within an otherwise specificational cleft.

For all of these authors then, we do not have a single *it*-cleft construction which permits both specificational and predicational readings. Instead we have an inherently specificational *it*-cleft and a separate, unrelated construction. In what follows, I explain some of the reasoning behind this claim. I show that the arguments in favour of a separate analysis for non-specificational examples are circular and originate from the (as I see it, incorrect) assumption that specificational *it*-clefts require an expletive analysis. I conclude that the lack of agreement between Ball (1977) and Declerck (1983) over how to analyse "borderline" non-proverbial predicational *it*-clefts calls for a unified analysis of all cleft types.

Although Ball (1977), Declerck (1983) and Lambrecht (2001) do not provide formal analyses of the grammatical structure of the *it*-cleft, they nevertheless adopt an expletive approach, in which the cleft clause is understood to be related in some way to the postcopular constituent. As I noted in §3.1, the expletive approach to *it*-clefts emphasizes the relationship between *it*-clefts and their noncopular subject-predicate counterparts. For example, the *it*-cleft in (7) is understood to express the same semantic content as its canonical counterpart in (8); that is, the clefted constituent is treated as the preposed argument of the embedded verb. For Ball, Declerck and particularly Lambrecht, *it*-clefts are therefore considered foremost as information-structure variants of noncopular sentences.

(7) It was [Howard]$_i$ that [Sarah kissed ____$_i$] [*it*-cleft]

(8) Sarah kissed Howard [canonical]

As I noted in §4.1.3, if the cleft clause is understood in relation to the postcopular constituent, as it is in the expletive accounts, it is shown to exhibit an interesting mixture of properties common to both restrictive and nonrestrictive relative clauses, thereby evading classification. For example, like restrictive relatives, the cleft clause is commonly introduced by *that* and can also occur without an overt relative

pronoun. However, the assumed antecedent to the cleft clause can be a proper name. Since proper names are full noun phrases, they cannot be further modified by restrictive relatives, shown in (11). This suggests that the relationship between the cleft clause and its immediate "antecedent" is different from that found with restrictive relative clauses in other constructions.

(9)	It was Howard that/Ø Sarah kissed	[*it*-cleft]
(10)	The guy that/Ø Sarah kissed works there	[restrictive relative]
(11)	*Howard that Sarah kissed works there	[restrictive relative]

In light of these unusual properties, these authors differ as to how they analyse the cleft clause. For instance, Lambrecht (2001: 473) argues that it is a special type of nonrestrictive relative, unique to cleft sentences, Declerck (1983:12) states that the *that*-clauses of clefts are not genuine relative clauses, whereas Ball (1977: 58) suggests that their internal structure is like a restrictive relative clause, but claims that they have "no apparent head". Despite this, what these authors all agree upon is that the cleft clause is not a normal restrictive relative clause.

However, unlike specificational *it*-clefts, proverbial examples do not correspond in meaning to their noncopular counterparts. Instead, these authors agree that they are paraphrased most closely by predicational copular sentences, such as (14).

(12)	It's a long road that has no turning	[proverbial cleft]
(13)	A long road has no turning	[canonical]
(14)	The/A road that has no turning is a long one/road	[predicational paraphrase]

This fact alone is enough for Lambrecht (2001: 503) to conclude that the proverbial examples have a "noncleft status". For him, the concept "cleft" is inextricably tied to the correspondence of these examples to simple noncopular sentences.

Proponents of the expletive analysis are therefore forced to examine specificational and proverbial examples in relation to two different constructions. As a

result, Ball, Declerck and Lambrecht are obliged to analyse the relative clauses of specificational and proverbial clefts differently. While in specificational *it*-clefts, the cleft clause is related in some way to the postcopular constituent, proverbial clefts cannot be accommodated into this analysis. Instead, Ball, Declerck and Lambrecht provide proverbial clefts with an analysis that corresponds to that of their predicational *NP be NP* paraphrases. Since sentences such as (14) above clearly involve restrictive relative clauses, which modify the head noun within the subject NP, these authors agree that the "cleft clause" of proverbial clefts is also a restrictive relative.

To recap then, Ball (1977), Declerck (1983) and Lambrecht (2001) examine specificational and proverbial clefts in relation to two very different constructions (noncopular simple sentences and predicational *NP be NP* sentences, respectively). As a result, they provide the sentence-final clauses of these cleft types with two different analyses: proverbial clefts are said to contain restrictive relative clauses while the cleft clause in specificational *it*-clefts is understood to be of a type that is unique to this construction.

In a circular line of argumentation, the status of the cleft clause is then used as evidence that the proverbial examples form a structurally distinct sentence type from specificational *it*-clefts. Declerck (1983: 14) says, "[proverbial clefts] are not really cleft sentences but represent a type of sentence that is homophonous with clefts. They differ from clefts in that they involve a true restrictive relative clause". However, as we have seen, these so-called differences in the status of the cleft clause are a consequence of relating specificational and proverbial clefts to two different constructions. Consequently, the status of the cleft clause cannot be used as a valid argument for treating proverbial clefts as a separate construction from specificational *it*-clefts.

There is also an important problem with setting up a sharp distinction between specificational *it*-clefts and proverbial "cleft-like" examples: that is, the status of non-proverbial predicational *it*-clefts. While Lambrecht (2001) does not discuss these examples, Ball (1977) and Declerck (1983) classify them with proverbial and specificational *it*-clefts, respectively. However, as I go on to explain in the remainder of

this section, these non-proverbial predicational clefts seem to bridge the gap between specificational and proverbial examples, thereby calling for a unified analysis.

For Ball (1977: 61), predicational clefts such as (15) should be analysed in the same way as the proverbial sentences, since they too correspond to predicational *NP be NP* sentences.

(15) It was a tall man that left [predicational cleft]

(16) The man that left was a tall man [predicational paraphrase]

However, there is less reason to assume that non-proverbial predicational examples are unrelated to the specificational *it*-cleft construction. For example, Declerck (1983: 15) notes that, like other proverbs, proverbial clefts are "reminiscent of an older stage of the language". He uses this claim to assert that these examples share only a superficial similarity to the modern-day specificational *it*-cleft. However, this argument cannot be made for non-proverbial predicational clefts. In addition, non-proverbial predicational *it*-clefts are often difficult to distinguish from their specificational counterparts, with the same example allowing for both specificational and predicational interpretations.

As a result, Declerck (1983) incorporates these non-proverbial predicational examples into his analysis of *it*-clefts. However, in order to maintain the claim that *it*-clefts are "essentially specificational in nature", he suggests that these examples "are really borderline cases" in which predicational information is contained within an otherwise specificational cleft (Declerck 1983: 18). For instance, in Declerck's account, the example in (17) involves a predicational adjective, but a specificational head noun. He notes that "If we leave out the predicational modifier, what remains must be a good specificational cleft" (Declerck 1983: 31).

(17) It was a charming woman that got the job ["borderline" cleft]

(18) It was a woman that got the job [specificational cleft]

However, Declerck acknowledges that for some examples the head noun cannot be given a specificational interpretation; here the entire focal noun phrase can be nothing other than predicational. For instance, in (19) and (20), the use of the degree modifier *such* and the negative marker *no* indicate that the head noun *idiot* is clearly a predicational element.

(19) It was such an idiot that wrote this

(20) It was no idiot that wrote this

For Declerck (1983: 38), these are treated as "exceptional" examples, in which "the essentially specificational nature of the cleft construction" has been "overruled". Nevertheless, the existence of true predicational *it*-clefts means that the claim that cleft sentences are inherently specificational cannot be maintained. The status of these "borderline cases" therefore poses a problem for accounts that seek to separate proverbial examples from true specificational clefts. If predicational *it*-clefts are possible, why should proverbial examples, which also have a predicational meaning, require a separate analysis?

In conclusion then, the claim that *it*-clefts are inherently specificational cannot be maintained. Instead, there are both specificational and predicational (including proverbial) *it*-clefts, which require a unified analysis. However, as we have seen, non-specificational examples cannot be successfully integrated into an expletive analysis of *it*-clefts. In such accounts, the clefted constituent is treated as the preposed argument of the cleft clause. Since predicational and proverbial examples cannot be interpreted in this way, I claim that this data poses a problem for the expletive approach. In other words, the fact that the expletive analysis cannot accommodate the full range of data raises the question of whether it should be employed to handle the subset of specificational *it*-clefts.

5.2.2 Integrating predicational and proverbial clefts into my analysis

In the previous section then, we have seen that expletive accounts of *it*-clefts are unable to provide a unified analysis which encompasses specificational, predicational and proverbial clefts. The reason for this is that, for these authors, the cleft clause in specificational *it*-clefts is understood to be related in some way to the postcopular element. As a result, this clause cannot be classified as a normal restrictive relative. From this, it follows that the specificational *it*-cleft construction has a different structure from non-specificational clefts, which are consistently analysed as containing restrictive relatives.

However, in §4.1, I argued in favour of a particular type of extraposition account of *it*-clefts, in which the constituent *it* and the cleft clause are treated together like a discontinuous definite description. One important advantage to this approach is that, by analysing the cleft clause as modifying the constituent *it*, rather than the complement of *be*, the cleft clause can be unambiguously classified as a restrictive relative. In this analysis, specificational *it*-clefts, like predicational and proverbial clefts, are understood in relation to corresponding copular paraphrases, shown in (21) and (22). As a result, all cleft types can be given a unified analysis. For instance, the predicational and proverbial clefts in (23) and (25) are likewise treated as involving discontinuous definite descriptions.

(21)	[It] was Howard [that left]	[specificational *it*-cleft]
(22)	[The one that left] was Howard	[specificational copular paraphrase]
(23)	[It] was a tall man [that left]	[predicational *it*-cleft]
(24)	[The man that left] was a tall man	[predicational copular paraphrase]
(25)	[It]'s a long road [that has no turning]	[proverbial *it*-cleft]
(26)	[The road that has no turning] is a long road	[predicational copular paraphrase]

In this analysis then, specificational *it*-clefts are treated foremost as specificational copular sentences, rather than as information structure variants of their noncopular counterparts. As Hedberg (1990, 2000) and Han and Hedberg (2008)

observe, the existence of specificational and predicational *it*-clefts becomes predictable in an account which views clefts as copular constructions, since noncleft *NP be NP* sentences also exhibit the specificational-predicational distinction. For instance, while (27) identifies the referent *Sam*, (28) ascribes a property to *the murderer*.

(27) The murderer is Sam [specificational copular sentence]
(28) The murderer is a doctor [predicational copular sentence]

This contrasts with Declerck's approach, in which predicational examples are effectively "hidden". For Declerck, only sentences that are overtly marked in a way that disallows a specificational interpretation are permitted into the "exceptional" class of true predicational clefts. All other predicational examples are integrated into a specificational analysis of *it*-clefts. For instance, Declerck analyses examples such as (18), repeated here as (29), as specificational, since they permit a reading in which the noun phrase (in this case *a woman*) is understood to denote a specific referent. Likewise, examples such as (30) are classified as "borderline" cases, which contain both predicational and specificational components.

(29) It was a woman that got the job [specificational cleft]
(30) It was a charming woman that got the job ["borderline" cleft]

A unified analysis, on the other hand recognizes that example (29) can also allow a predicational reading where the postcopular indefinite noun phrase *a woman* is understood to describe a class or kind rather than referring to an individual (see Hedberg 1990: 55). On this purely predicational reading, the gender of the referent is at issue, rather than her identity. Although this sentence does not contain any disambiguating markers, the very fact that it **can** (shown in (31)) indicates that a predicational reading is possible for the unmodified postcopular NP in (29).

(31) It was certainly a woman that got the job [predicational cleft]

Also, by accepting that *it*-clefts are not inherently specificational, we do not have to classify examples such as (30) above as "borderline" cases. Given the analysis of specificational meaning outlined in §3.3, it does not make sense to speak of specificational and predicational "components", since specificational meaning is the result of a special type of nominal predication relation. Instead, examples such as (30) are acknowledged to have a fully predicational reading, corresponding to noncleft predicational copular sentences such as (32).

(32) The woman that got the job was a charming woman

Consequently, an important problem with Declerck's (1983) approach is that by incorporating predicational examples into an essentially specificational analysis of *it*-clefts, the specificational-predicational distinction is dismissed rather than explained. In contrast, using the discontinuous constituent analysis advocated here, we are able to recognize the distinction in meaning between specificational and predicational (including proverbial) clefts, while at the same time integrating them into a unified account of the *it*-cleft construction.

The next step then is to ask how these different interpretations for *it*-clefts come about. Crucially, this depends on how we analyse the distinction between specificational and predicational copular constructions. In what follows, I show that the relationship between specificational and predicational meaning as it is outlined in §3.3 is particularly good at both explaining why *it*-clefts exhibit the specificational-predicational distinction and accounting for the predicational *it*-cleft data.

In §3.3, I argued that specificational meaning arises from a reinterpretation of the nominal predication relation associated with definite noun phrase predicates. In §4.1, I integrated specificational *it*-clefts into my analysis of specificational copular constructions, arguing that the constituent *it* and the relative clause function as the definite NP predicate. In this analysis, specificational *it*-clefts, such as (33), and their corresponding copular paraphrases, such as (34), involve a non-referring definite description and a postcopular referential NP. These noun phrases enter into a class-

membership predication relation, in which the referent is understood to be a member of a set. In such constructions, the set described by the definite NP is so restricted that the referent makes up the complete membership of this set. Consequently, a specificational meaning ensues, whereby the referent, which is in focus, is identified as the sole member of the restricted set.

(33) It was John that won the prize

(34) The one that won the prize was John

In order to provide an analysis for predicational *it*-clefts, we first need to return to the semantics of definite noun phrases detailed in §3.3.2. From a Russellian perspective, definite noun phrases are semantically descriptive, non-referring expressions. However, definite noun phrases can of course be used referentially, to pick out a particular individual. According to Grice (1969), this usage is a pragmatic, rather than a semantic phenomenon, involving the exploitation of speaker and hearer assumptions.

In predicational *it*-clefts, we have the same basic composition as specificational *it*-clefts, which comprise of a discontinuous definite description and a postcopular NP. However, these elements perform different roles in the predication relation. For example, in the predicational *it*-cleft in (35) and its paraphrase in (36), the definite description *it (the one) that won the prize* is used referentially. In other words, the Gricean reasoning outlined above has gone through. In contrast, the postcopular element *a woman* is understood to denote a class or kind. These two noun phrases enter into a nominal predication relationship, such that the referent *the one that won the prize* is understood to be a member of the general set *woman*. In such constructions, the predicate nominal is located in the postcopular (focal) position. In (35) and (36) then, the property of being *a woman* is attributed to *the one that won the prize*.

(35) It was a woman that won the prize

(36) The one that won the prize was a woman

This same analysis can also be extended to proverbial clefts. In such cases, the definite description is used to refer to a hypothetical entity. For example, in the *it*-cleft in (37) and its predicational copular paraphrase in (38), the mental referent *the mother that has such children* is described as *a happy mother*.

(37) It is a happy mother that has such children

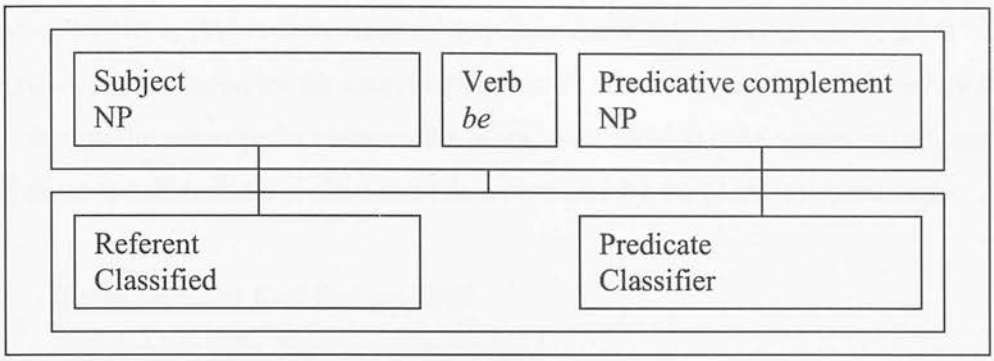
(38) The mother that has such children is a happy mother

In predicational and proverbial *it*-clefts, the postcopular element is always nominal.[2] However, in many cases, the postcopular NP contains a premodifying adjective. For instance, in (37) above, the adjective *happy* bears the main focal stress, while the noun *mother* is given information. In such examples, the head noun of the postcopular NP provides us with a value for the underspecified head noun contained within the constituent *it*. For instance, in the *NP be NP* paraphrase in (38), the head noun (or "background set") of the initial definite description matches the head noun of the postcopular NP.

I suggest that predicational *it*-clefts, of which the proverbial examples form a subtype, inherit from the predicate nominal construction outlined in §3.3.3. While specificational *it*-clefts inherit the *subject-be-complement* syntactic structure of the (canonical) specificational copular construction, non-specificational examples inherit the *subject-be-predicative complement* structure of the predicate nominal construction (see Figure 5.1). This accounts for the fact that indefinite noun phrases can have a predicative function in *it*-clefts such as (35) and (37) above. As I explained §3.3.3, the predicate nominal construction provides indefinite noun phrases with a predicative semantics via coercion (or construal). In contrast, indefinite noun phrases cannot be readily interpreted as predicates in the specificational copular construction unless they share characteristics with definite noun phrases, which are inherently set-denoting (see §3.3.4).

[2] This fact alone provides support for my analysis of specificational *it*-clefts as involving a type of nominal predication relation; it shows that all types of *it*-cleft (both specificational and predicational) involve nominal (as opposed to any other type of) predication.

The predicate nominal construction



The predicational it-*cleft construction*

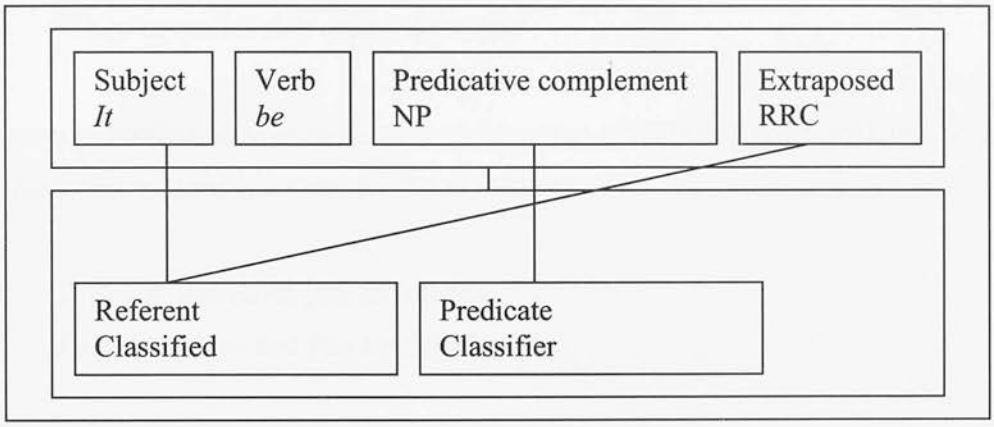


Figure 5.1 *The predicational* it-*cleft construction and the overarching predicate nominal schema*

This unified account of *it*-clefts analyses both specificational and predicational (including proverbial) *it*-clefts as involving nominal predication. For specificational *it*-clefts, the discontinuous definite description functions as the semantic predicate, while in predicational *it*-clefts this discontinuous constituent is treated as an act of reference. The specificational-predicational distinction in *it*-clefts therefore falls out from an analysis in which *it* and the relative clause are treated as a discontinuous definite description. Since definite noun phrases are semantically descriptive, but nevertheless allow a referential usage, we would expect that both readings are possible for the discontinuous constituent in *it*-clefts.

My analysis is further supported by differences in the specificational and predicational *it*-cleft data. As pointed out by Ball (1977), specificational *it*-clefts with

plural foci are always introduced by a singular cleft pronoun. For example, while the postcopular NP in (39) is comprised of two individuals *John and Karl*, the cleft is nevertheless introduced by the singular pronoun *it*. Likewise, (41) is introduced by *it* even though the postcopular element is a plural noun (*wolves*). As shown in (40) and (42), these specificational clefts cannot be introduced by the plural pronoun *they*.

(39) It was John and Karl that got fired

(40) *They were John and Karl that got fired

(41) It's wolves that are my favourite

(42) *They're wolves that are my favourite[3]

In contrast, predicational *it*-clefts with plural nouns in the postcopular position can be introduced by a plural pronoun, shown in (43) and (44).

(43) They are nice shoes you're wearing

(44) They're wolves that you can hear howling

This data indicates a semantic difference in the discontinuous definite descriptions of these two different types of cleft. Consequently, these examples provide support for the account of *it*-clefts presented in this thesis, in which the definite-like NP of specificational *it*-clefts is analysed as a non-referring set description while that of predicational *it*-clefts acts as a referring expression.

Furthermore, my analysis goes some way to explaining this data. I have argued that the discontinuous definite description of specificational *it*-clefts always denotes a set. This property is shared with other noncleft *NP be NP* specificational sentences

[3] Example (42) cannot be given a predicational reading because we cannot interpret *my favourite things* as a referring expression in this context. For instance, (i) below will always be read as a specificational sentence, with *my favourite things* acting as the definite NP predicate. This specificational interpretation makes much more sense than the predicational reading, which describes all of my favourite things as wolves.

(i) My favourite things are wolves

containing definite NP predicates. However, in §4.3.2, I noted that specificational *it*-clefts behave differently from specificational sentences with overt nominal heads. For example, both the *NP be NP* sentence (or *th*-cleft) in (45) and the *it*-cleft in (46) contain plural foci in the postcopular position. While the subject NP in (45) contains a nominal head *ones* which is also marked as plural, the constituent *it* and the copular verb of (46) are marked as singular.

(45) The **ones** that travel by train **are John, Suzie and Tom**

(46) **It's John, Suzie and Tom** that travel by train

In §4.3.2, I suggested that a plausible explanation for these number agreement patterns lies in the analysis of the cleft *it* as performing the functions of both the definite article and the head noun of the description. Since this "head noun" is underspecified, the discontinuous definite description denotes a singular set, but does not mark the membership of this set as either singular or plural. I suggested that this property of *it*-clefts could play an important role in manufacturing the interesting number agreement patterns which are found in *it*-clefts but are not shared by other kinds of specificational copular construction. I concluded that in specificational *it*-clefts with plural foci, such as (46) above, a singular set (*of co-workers that travel by train*) is identified as having plural members (*John, Suzie and Tom*). From this, it follows that the definite-like description in *it*-clefts will always be introduced by a singular pronoun, since it denotes a singular set without indicating the plurality of its members.

In specificational *it*-clefts then, the discontinuous definite description denotes a singular **set**. However, in predicational *it*-clefts, the discontinuous definite description is used to refer to the individual or **individuals** that are the member(s) of some set. In this analysis, the use of the plural pronoun *they* to introduce predicational clefts with plural foci is therefore expected.

In conclusion, while the predicational *it*-cleft data is at odds with the expletive approach to *it*-clefts (which cannot provide a unified analysis of specificational, predicational and proverbial clefts), I have shown that it can be successfully

accommodated by, and therefore provides support for, the analysis of *it*-clefts presented in this thesis. Here, *it*-clefts are understood in relation to specificational and predicational *NP be NP* sentences, both of which involve nominal predication. As a result, both predicational and specificational *it*-clefts are analysed as involving a discontinuous definite description which enters into a class-membership predication relation with the postcopular NP. The interpretation of this discontinuous constituent as either a referring expression or as a description of a set therefore plays a crucial role in manufacturing the predicational and specificational distinction in *it*-clefts.

However, this is not to say that all "predicational" accounts of specificational copular sentences are equally successful at explaining the relationship between predicational and specificational *it*-clefts. For example, Heggie (1988) and Mikkelsen (2005) both assume inverse analyses of specificational *NP be NP* sentences; that is, they argue that such sentences are derived via a movement operation resulting in the predicative NP preceding the underlying subject NP (see §3.2.2). However, when it comes to specificational *it*-clefts, neither author recognizes that these too involve nominal predication. Instead, they suggest that the cleft clause is predicated of the clefted constituent, with the initial *it* either acting as a semantically dummy element (Heggie 1988) or as a property-denoting pronoun which is bound cataphorically to the cleft clause (Mikkelsen 2005: 121).

While neither of these authors discuss *it*-clefts with a predicational meaning, it is difficult to see how their analyses could account for them. If we take the postcopular NP of these *it*-clefts to be the semantic predicate, then what is it predicated of? How can the sentence-final clause be interpreted as a referring expression? Without analysing the initial *it* and the cleft clause as forming a discontinuous definite NP, these authors cannot integrate *it*-clefts into a unified analysis of *NP be NP* sentences.[4] As a result, they do not predict that *it*-clefts will display the specificational-predicational distinction.[5]

---

[4] I discuss the problems with Heggie's (1988) account of *it*-clefts in §3.1 and in §4.1.1. Although Heggie assumes an inverse analysis of specificational *NP be NP* sentences, she nevertheless proposes an expletive analysis of *it*-clefts.

[5] However, it does not follow that the inverse approach to specificational copular sentences cannot be extended to account for the specificational-predicational distinction in *it*-clefts. For example, Den Dikken (2009) assumes that while the initial *it* of specificational *it*-clefts is a (meaningless) pro-predicate (which

In conclusion then, the particular analysis of *it*-clefts presented in this thesis is able to both accommodate *it*-clefts containing predicative foci and explain their relationship to specificational clefts. In this analysis, *it*-clefts can be reduced to *NP be NP* sentences containing definite noun phrases. Just as noun phrases can function both as referring expressions and as predicates, so to can the focal NP and the definite-like description in *it*-clefts. The analysis presented in this thesis is therefore an improvement on the expletive analyses (see §5.2.1) and on other accounts that do not integrate *it*-clefts into an analysis of *NP be NP* sentences (including Heggie 1988 and Mikkelsen 2005).

5.2.3 Specification and predication or equation and predication?

In this section, I compare my analysis to that of Han and Hedberg (2008), who also extend their account to accommodate predicational *it*-clefts. Like me, these authors suggest that both specificational and predicational *it*-clefts contain a discontinuous definite description. However, Han and Hedberg (2008) advocate an equative, rather than a predicational, analysis of specificational copular sentences, arguing that both the discontinuous constituent and the focal element are semantically referring. Given that both Han and Hedberg's (2008) and my analysis have a number of things in common, this comparison allows me to focus on comparing an equative account of specificational sentences to a predicational one. I use this opportunity to raise further objections to the equative approach outlined in §3.2.1, and go on to show that my predicational analysis of specificational meaning is better at explaining the behaviour of specificational and predicational *it*-clefts. I conclude that my understanding of specificational meaning as involving a nominal predication relation results in a more economical analysis for the *it*-cleft than that proposed by Han and Hedberg (2008).

---

has undergone inversion), the *it* of predicational *it*-clefts is a referential pronoun. Unlike the analysis proposed in this thesis, Den Dikken argues that the relative clause in specificational *it*-clefts is not a restrictive relative, but a (base-generated) right-dislocated constituent containing a null-head which enters into a concord relationship with the focal element. Although Den Dikken does not explicitly address the nature of the relative clause of predicational *it*-clefts, such examples could be accommodated into this analysis. Nevertheless, I reject Den Dikken's (2009) account for two reasons. First, the positing of phonologically null elements is against the principles that underlie a constructional analysis and secondly, I believe that there is more evidence in favour of a restrictive relative analysis of the cleft clause than a free relative analysis (see §4.2).

Like me, Han and Hedberg (2008) argue that *it*-clefts should be analysed in relation to *NP be NP* sentences. However, while I claim that specificational copular sentences involve a nominal predication relation, Han and Hedberg adopt Heycock and Kroch's (1999) analysis of copular constructions. These authors assume that there are only two types of copular sentence: equative and predicational. As a result, all types of specificational sentence are analysed as equative structures, involving two arguments of the same semantic type.

For Han and Hedberg (2008) then, both the *it*-cleft in (47a) and its ordinary copular paraphrase in (47b) contain a two-place equative predicate. Since the proper name *Ohno* is a referring expression, the (discontinuous) definite description, *it (the one) who won*, must also be treated as referential in order for them to be equated.

(47) a. It was Ohno who won
b. The one who won was Ohno
c. THEz [won(z)] [z = Ohno] (Han and Hedberg 2008: 349)

In contrast, predicational *it*-clefts, such as (48a), and their corresponding copular sentences, such as (48b), do not involve noun phrases of the same semantic type. Here, the (discontinuous) definite description is again treated as a referring expression, but the post-copular element is property-denoting. These constructions therefore involve a predication relationship, ascribing a property to the definite noun phrase referent.

(48) a. It was a kid who beat John
b. The one who beat John was a kid
c. THEz [beat(z, John)] [kid(z)] (Han and Hedberg 2008: 349)

While Han and Hedberg's (2008) semantic analysis of predicational *it*-clefts is therefore the same as that argued for in §5.2.2, their analysis of specificational *it*-clefts is very different. In what follows, I show that the relationship between predicational and specificational *it*-clefts cannot be satisfactorily explained by Han and Hedberg's (2008)

analysis. However, first, I argue that the analysis of copular constructions employed by Han and Hedberg (and outlined originally by Heycock and Kroch (1999)) is fundamentally flawed.

Heycock and Kroch (1999: 374) argue that analysing specificational sentences as semantically equative has an "economy advantage", since examples such as (49), which involve two equally referential arguments, do not require a separate analysis. For these authors, the taxonomy of copular sentences is therefore reduced to just two types: equative and predicational.

(49) Cicero is Tully [equative]

(50) Cicero is an orator [predicational]

However, in order to accommodate specificational copular sentences into an equative analysis, Heycock and Kroch (1999) provide them with a semantic interpretation which, they recognize, is counterintuitive. For instance, in discussing the *wh*-cleft presented here as (51), Heycock and Kroch (1999: 383) note that although *what Fiona bought* seems to denote a **set** of individuals, "a set cannot be directly equated with an individual". As a result, they conclude that this free relative actually denotes "the maximal individual" or "a plural individual".

(51) What Fiona bought was that ancient dictionary

Heycock and Kroch (1999: 373) therefore recognize that the initial NP of specificational copular sentences is frequently "less referential" than the postcopular NP; an observation supported by Mikkelsen's (2005) pronominalization results (see §3.2.2). Nevertheless, this asymmetry is explained away and the data is coerced to fit in with the equative analysis. Aside from the economy advantage, it is difficult to find justification for this. For instance, if economy is our only goal, then it must be equally acceptable to argue that there are in fact no true equative sentences and that all copular constructions involve an asymmetric predication relation.

Indeed, this is what Partee (1986), a proponent of the inverse analysis of specificational sentences, suggests. Formalizing the semantic analysis proposed by Williams (1983), Partee (1986) argues that *be* takes two arguments, one referential (type *e*) and one predicative (type <*e,t*>), which can appear in either order. Just as Heycock and Kroch make use of semantic operators to obtain a referential reading for the initial NP, Partee (1986: 356) draws on type-shifting principles in order to accommodate "apparent identity statements" into her analysis.[6] For example, Partee makes use of the 'ident' operation to get a predicative reading from the referential proper name *Mr. Smith* in the example given here as (52). This operation converts a type *e* expression into a type <*e,t*> expression, therefore enabling (52) to be reanalysed as a predicational sentence.

(52) John is Mr. Smith (Partee 1986: 356)
e <*e,t*>

Despite their different approaches then, both Heycock and Kroch (1999) and Partee (1986) advocate unified analyses of specificational and equative copular sentences. However, while this may result in a more economical taxonomy, I believe that it is a false economy. For one thing, it is not supported by our intuitions. As Heycock and Kroch (1999: 381) acknowledge, specificational copular sentences "are asymmetric in interpretation in a way that "true equatives" are not". Secondly, by subsuming equative and specificational sentences under the same analysis, we end up sacrificing economy elsewhere in the system.

For example, like Partee (1986), Den Dikken (2006) argues that there are only two types of copular sentence, canonical and inverse, both of which involve predication. While Den Dikken recognizes that identity statements are semantically equative, he

[6] A problem with invoking type-shifting principles is that while they inform us that the semantic denotation has changed, they do not provide an explanation as to why speakers type-shift. For this reason, I have made use of Gricean reasoning to explain how a semantically non-referring definite NP can, in many constructions, be used as a referring expression. This pragmatic explanation prioritizes the role of the interlocutors and is supported by Mikkelsen's (2005) pronominalization data. Similarly, in §5.5, I argue that type-shifting does not provide an adequate explanation for why *it*-clefts permit non-NP foci. Consequently, in later chapters, I make use of the more explanatory concept of 'coercion'.

nevertheless categorizes them with "inverse" specificational sentences, arguing that they have an asymmetrical structure.[7] However, in order to provide equatives like (53) with a unified asymmetrical syntax, Den Dikken (2006: 72) has to posit an enormous amount of phonologically empty structure, involving three separate stages.

First, he must assume that equative sentences contain a small clause structure (labelled *RP* for Den Dikken) with a phonologically null 'pro-predicate', which is associated with a reduced relative clause. Together, these components form a free relative. This is shown in (54), in which the underlying structure of *Cicero* in (53) is analysed as a free relative, much like *what Cicero is*. The reduced free relative *Cicero* is then predicated of the subject *Tully* as part of another small clause structure, shown in (55). Finally, the predicate undergoes inversion and is raised to SpecTP, resulting in a uniform analysis with specificational copular constructions.

(53) Cicero is Tully

(54) [PRO-PREDICATE$_{\varnothing}$ [$_{CP}$ *Op*$_i$ [C$_{\varnothing}$ [$_{RP}$ *Cicero* [RELATOR$_{\varnothing}$ *t*$_i$]]]]]

(55) [$_{RP}$ *Tully* [RELATOR [$_{Pred}$ PRO-PREDICATE$_{\varnothing}$ [$_{CP}$ *Op*$_i$ [C$_{\varnothing}$ [$_{RP}$ *Cicero* [RELATOR$_{\varnothing}$ *t*$_i$]]]]]]]

(56) [$_{TP}$ [$_{Pred}$ PRO-PREDICATE$_{\varnothing}$ [$_{CP}$ *Op*$_i$ [C$_{\varnothing}$ [$_{RP}$ *Cicero* [RELATOR$_{\varnothing}$ *t*$_i$]]]]]$_j$ [T+ RELATOR$_k$ =*be* [$_{RP}$ *Tully* [t$_k$ t$_j$]] (Den Dikken 2006: 72-73)

Consequently, in order to achieve a unified analysis of equative and specificational sentences, Den Dikken provides equatives with a complex asymmetric structure which

[7] Like Den Dikken (2006), Heggie (1988) also claims that while equative sentences such as (i) contain two referring expressions, all types of copular sentence involve the same subject-predicate configuration at D-structure. In order to make this work, Heggie treats equatives as "naming" sentences. For example, she claims that the postcopular proper name in (i) is a syntactic predicate and is therefore "treated as a label" (Heggie 1988: 10). However, Heggie conflates two very different readings of (i) in her account. On the equative reading, *that man over there* is identified as the same person as the individual we know as *Jack Jones*. On the "identificational" reading, paraphrased in (ii), we are told the name of *that man over there*; it is only on this reading that *Jack Jones* behaves like a label or naming device (see Higgins 1979).

(i) That man over there is Jack Jones (Heggie 1988: 98)
(ii) That man over there is called Jack Jones

seems unmotivated given their simple surface form. As a result, the advantages of an economical system are countered by a complicated underlying structure.[8]

Likewise, in what follows, I show that the equative analysis of specificational sentences adopted by Han and Hedberg (2008) is unable to provide an explanatory and economical account of the relationship between specificational and predicational *it*-clefts. As I indicated above, Han and Hedberg (2008) assume that the discontinuous definite description of both specificational and predicational *it*-clefts is always referential. The interpretation of an *it*-cleft as specificational (their equative) or predicational therefore depends solely upon whether the postcopular element is understood to be a referring or a property-denoting expression. However, structurally, "equative" and predicational *it*-clefts are very different. Han and Hedberg (2008) assume that equative and predicational *it*-clefts contain two different forms of the copular verb; while the "equative copula" is a predicate taking two arguments, the "predicational copular" tree involves a nominal predicate and a single argument.

There are several problems with this analysis which are not shared by the account proposed in §5.2.2. First, Han and Hedberg (2008) cannot explain Ball's (1977) observation that only predicational clefts can be introduced by the plural pronoun *they*. They note, "Why equative clefts require singular cleft pronouns when they contain a plural clefted constituent does not follow from our theory and remains a puzzle" (Han and Hedberg 2008: 372). As I noted above, this difference in behaviour between specificational and predicational clefts indicates a semantic difference in the discontinuous definite descriptions of these two interpretations. This is not present in the analysis of Han and Hedberg, who argue that the discontinuous constituents of both equative and predicational clefts are referential.

[8] In §3.3.3, I claimed that equative and specificational sentences have a different semantics but a common syntactic structure. The constructional framework I adopt does not require me to have a complex syntactic architecture in order to establish maximum isomorphism between syntax and semantics. In my analysis, the predication relation in specificational sentences is derived compositionally from the semantics. Unlike Den Dikken's (2006) minimalist account, I can therefore analyse specificational and equative constructions as having the same structure without assuming that the latter is underlyingly predicational. Consequently, my analysis does not reduce economy elsewhere in the system.

As well as finding it difficult to accommodate and explain the full range of data, Han and Hedberg's (2008) analysis is less economical than that proposed in §5.2.2. As noted above, Han and Hedberg claim that equative and predicational clefts are structurally very different, involving two distinct kinds of copular verb. In contrast, in the analysis outlined in §5.2.2, both specificational and predicational *it*-clefts involve nominal predication. Consequently, there is no need to assume that these different types of cleft involve different forms of the copular verb or radically different structures. Instead, the distinction between specificational and predicational meaning in *it*-clefts depends crucially on the interpretation of the discontinuous definite description (*it* and the restrictive relative clause). In turn, this rests on the Russellian-Gricean relationship between the non-referring semantics and referential usage of definite noun phrases.

In conclusion then, predicational *it*-clefts can be more successfully integrated into an account of copular *NP be NP* sentences that analyses specificational meaning as involving nominal predication. The predicational *it*-cleft data therefore provides further evidence against an equative analysis of specificational copular sentences.

Over the last three sections, I have shown that a comprehensive analysis of predicational and proverbial clefts falls out from the accounts of specificational meaning, definite noun phrases, and *it*-clefts developed throughout this thesis. I have argued that while predicational *it*-clefts have much in common with specificational *it*-clefts, they inherit from the predicate nominal construction rather than the (canonical) specificational copular construction. However, not all types of (specificational) *it*-cleft can be dealt with by inheritance from the language system.[9] In the following two sections, I discuss *it*-clefts with non-NP foci and 'informative presupposition' *it*-clefts, respectively. As I explain, the range of *it*-cleft foci and the discourse status of the cleft clause are construction-specific properties, requiring an idiosyncratic description. I go on to show that they can only be explained by looking for motivation, not within the language system, but in the construction's development over time.

---

[9] In keeping with the cleft literature, I continue to use the term '*it*-cleft' as a convenient shorthand for 'specificational *it*-cleft'. In the remainder of this thesis, I explicitly indicate where the discussion turns to predicational *it*-clefts or relates to the overarching *it*-cleft schema which encompasses both of these subtypes.

5.3 Non-NP focus *it*-clefts

Most authors writing about *it*-clefts centre their discussion on examples containing noun phrases in the postcopular position. This is understandable, since *it*-clefts with non-NP foci are particularly puzzling and often pose a problem for analyses that seem to work well with the NP-focus *it*-cleft data. Fortunately, as I go on to show, non-NP *it*-clefts can be integrated fairly neatly into the analysis proposed in this thesis. Nevertheless, describing and explaining the exact range of elements which make up the list of possible *it*-cleft foci is not straightforward from the analysis developed so far. As I explain, this can only be achieved if we take historical evidence into consideration, a proposal which I follow up in chapter 7 of this thesis.

Many of the *it*-cleft analyses proposed in the literature cannot accommodate the full range of *it*-cleft foci. For example, in §4.2.5, I showed that extraposition accounts framed within a transformational theory of grammar (including Akmajian 1970, Gundel 1977 and Wirth 1978) assume that *it*-clefts should allow the same range of focal elements as their pseudocleft sources (or at least a proper subset of them). However, not all *it*-cleft examples have pseudocleft counterparts. For example, certain prepositional phrases will only be permitted as the focal element of an *it*-cleft, as shown in (57).

(57) a) It is with great pride that I accept this award
b) *How I accept this award is with great pride (Gundel 1977: 548)

As I explained in §4.2.5, these examples cannot be successfully accounted for in an extraposition analysis formulated within the transformational tradition.[10] Consequently, in order to account for the full range of *it*-cleft foci, Pinkham and Hankamer (1975) argue that *it*-clefts with non-NP foci are derivationally distinct from those with noun phrases in the postcopular position. They claim that only NP-focus *it*-clefts can be derived by extraposition (forming "deep-clefts"); *it*-clefts with non-nominal clefted

[10] Not without positing an unlikely source structure and extremely stipulative transformations (see Emonds 1976 and the discussion in footnote 17 of §4.2.5).

elements are derived from their corresponding noncopular sentences via extraction of the focus element (forming "shallow-clefts").

Expletive *it*-cleft analyses also have difficulties with incorporating the full range of possible focal elements. For example, as I explained in §3.1, Chomsky (1977) claims that the focal element is coindexed with a gap, generated by *wh*-movement, in the cleft clause. However, this analysis pertains only to *it*-clefts with NP and PP foci, which can (sometimes) be treated as preposed arguments of the verb embedded in the cleft clause. As a result, Chomsky suggests that *it*-clefts with adverbial foci require a separate syntactic derivation. Likewise, Delin (1989) also provides two different syntactic analyses for the *it*-cleft. She claims that the cleft clause in *it*-clefts with nominal foci is a restrictive relative. However, since restrictive relative clauses cannot modify non-nominal antecedents, she argues that non-NP focus *it*-clefts instead contain a "sentential complement". For Delin then, *it*-clefts with nominal and non-NP foci are structurally distinct (see §4.1.3).

However, there is no real evidence to suggest that *it*-clefts with different types of focal element should require a separate analysis[11]. Consequently, as I see it, an important advantage with my own analysis of *it*-clefts is that it can accommodate the full range of *it*-cleft foci. As I explained in §4.1.3, I assume that the cleft clause is connected to the initial *it*, rather than to the complement of *be*. With this analysis in place, the cleft clause can be consistently analysed as a restrictive relative regardless of the category of the clefted constituent, *contra* Delin (1989). As a result, *it*-clefts with all categories of foci are shown to make up a single construction.

Furthermore, unlike the transformational analyses of Akmajian (1970), Gundel (1977) and Wirth (1978), the particular extraposition analysis put forward in this thesis is non-derivational. As a result, a lack of isomorphism between *it*-clefts and other specificational copular constructions is not problematic for this account. As I explained in §4.2.5, the construction grammar framework I adopt tolerates and anticipates idiosyncratic, construction-specific information. In this model, constructions are

[11] See for example Gundel (1977), who disagrees with Pinkham and Hankamer's (1975) judgments regarding the relative acceptability of NP and non-NP focus *it*-clefts in environments involving extraction.

recognized as independent linguistic objects rather than epiphenomena. This, together with default inheritance (which allows overrides), means that on a constructional account, *it*-clefts can exhibit properties not shared by other subtypes of copular construction, despite the fact that they all inherit from the same basic schema. I claim that the range of *it*-cleft foci is a property specific to this construction, which therefore requires an independent explanation

It-clefts with all categories of focal element can therefore be accommodated straightforwardly in my constructional account. However, as yet, I have not put any restrictions in place which limit the range of focal elements to produce only grammatical *it*-cleft examples. Many of the analyses put forward in the cleft literature restrict the range of *it*-cleft foci by syntactic means. For example, as I noted in §4.2.5, Akmajian (1970) and Wirth (1978) suggest that their transformational rules of extraposition apply only to structures with NP (and, for Akmajian, PP) foci. However, on this account, as for many others, the range of *it*-cleft foci is grossly underestimated.

In contrast, Delahunty (1982, 1984) makes it his priority to present an analysis which accommodates the full range of *it*-cleft foci (including noun phrases, prepositional phrases, adverbial phrases, particles, adjective phrases and quantifier phrases) and excludes all other constituents from occurring in the postcopular position. As I noted in §4.3.1, Delahunty claims that the clefted constituent and the cleft clause are syntactic sisters within VP. Using Jackendoff's (1977) rules for the projection of V, Delahunty is able to sanction the full range of possible *it*-cleft foci while at the same time explaining why certain syntactic categories are excluded from occurring in the postcopular position. For example, Delahunty argues that the nonexistence of phrase structure rules such as VP → V VP S' accounts for the absence of *it*-clefts with VP foci, shown in (58).

(58) *[$_{S}$ It [$_{VP}$ [$_{V}$ is] [$_{VP}$ swum the ocean] [$_{S'}$ that Michael has]]]

However, in many ways, Delahunty's phrase structure rules are too general. Since these rules apply wholesale to each of the phrasal categories outlined above, they do not predict that some members of these categories will be more acceptable as cleft

foci than others. For instance, while both (59) and (60) are examples of adverb-focus *it*-clefts, they nevertheless differ in acceptability.

(59) ?It was slowly that Mary dressed to go out

(60) It wasn't only slowly that Mary dressed, but carefully too (Delahunty 1984: 80)

In order to explain why some members of a syntactic category are less suited to occurring in the focal position than others, Delahunty appeals to the *it*-cleft's pragmatic properties. He notes that example (59) is not as acceptable as (60) because the latter enhances the *it*-cleft's uniqueness and exhaustiveness implicatures through negation and the use of a focusing adverb (Delahunty 1984: 80).[12]

This begs the question, if we ultimately need to provide a pragmatic and/or semantic explanation for limiting the range of possible foci, what do we need the syntactic phrase structure rules for? From this, it seems that the range of *it*-cleft foci cannot be satisfactorily defined in terms of general syntactic rules or major syntactic categories.

---

[12] In addition, Delahunty (1984: 73) notes that his phrase structure rules do not explain why *it*-clefts only very rarely permit multiple foci, as shown in (i). Delahunty attempts to account for this by appealing to the *it*-cleft's pragmatic properties. He notes, "Because the focal constituent is that constituent of which the clause is uniquely and exhaustively true…there may be only one such constituent" (Delahunty 1984: 73). However, this generalization cannot explain attested counterexamples such as (ii).

(i) *It was Tom on Saturday that stole the money

(ii) It was at Knock a century ago that the Virgin appeared to local peasants (Delahunty 1984: 73)

I argue that, once we analyse the *it*-cleft as a subtype of specificational copular sentence, this data can be given a simple explanation. As shown in (iii) and (iv), the *it*-clefts above correspond to *NP be NP* sentences which show the exact same level of acceptability.

(iii) *The **person** and **day** that stole the money was Tom on Saturday

(iv) The **place** and **time** that the Virgin appeared to local peasants was at Knock a century ago

This suggests that, *contra* Delahunty (1984), it is the semantics of the head noun in the definite description, rather than the *it*-cleft's pragmatic properties, which govern the occurrence of multiple foci. While the concepts *place* and *time* are both adverbial nouns, requiring a restrictive relative clause without an overt gap, *person* and *day* cannot be coordinated within the same NP, since only the former requires a subject gap in the relative clause. The fact that the distribution of multiple foci in *it*-clefts patterns with the *NP be NP* data supports the main claims of this thesis: that the *it*-cleft is a subspecies of specificational copular sentence containing a discontinuous definite description with an underspecified head noun.

So how should we describe and explain the range of *it*-cleft foci? One option is to examine the semantic function of the focal element in cleft sentences. As a specificational copular construction, the *it*-cleft involves nominal predication (see §3.3.3). That is, the (discontinuous) definite NP is predicated of the referring expression contained within the postcopular slot. It is plausible then that the range of *it*-cleft foci is governed by semantic-pragmatic criteria: we would expect that only those elements that can be given a referential interpretation can occur in the postcopular, focal position. Of course, this approach cannot explain why *it*-clefts have a different range of foci from other specificational copular constructions. Nevertheless, it helps us to understand why some items are more likely to occur in the focal position than others.

One advocate of this approach is Partee (1986). She claims that since non-NP items typically denote properties, rather than objects, they are not well-suited to the referring function. In order to accommodate specificational sentences with non-NP foci, such as the *wh*-cleft in (61), Partee makes use of type-shifting principles. She claims that, in this example, the adjective *unusual* is nominalized via the type-shifting operation 'nom'. This operation "maps a (predicative) property on to its individual correlate", converting *unusual* from a type $<e,t>$ expression into a type *e* referring expression (Partee 1986: 357).

(61) What John is is unusual (Partee 1986:363)

Davidse (2000) provides a similar analysis, which focuses on *it*-clefts. She claims that the postcopular position is a "strictly 'nominal' slot"; all of the non-NP elements that occur in the postcopular position must be "rankshifted" into this category via nominalization (Davidse 2000: 1116).

However, I am skeptical of the type-shifting and rank-shifting approaches of Partee (1986) and Davidse (2000). While I agree that the postcopular item in *it*-clefts is a referring expression, I see no evidence that non-nominal *it*-cleft foci are "reclassified" as

noun phrases.[13] For instance, Partee (1986) observes that, in cleft sentences, non-NP foci do not display the same morphological markings as other nominalized forms. Discussing the *wh*-cleft given in (61) above, Partee (1986: 363) asks "why is the form here *unusual* rather than *unusualness*?" While Partee leaves this question unanswered, the data suggest that nominalization is not involved in sanctioning non-NP focus *it*-clefts. Instead, it appears that while the meaning of the non-NP element remains the same, its function is altered, and it is in this way that the item becomes a referring expression. I return to the construal of non-nominal *it*-cleft foci in chapter 7, where I invoke the concept of 'coercion'.

A better explanation for the range of *it*-cleft foci is provided by É. Kiss (1998). She argues that the focal element in "identificational" (our specificational) sentences must denote an individual. Since verb phrases and predicative elements do not denote individuals, É. Kiss' account explains why these items do not make acceptable *it*-cleft foci. Nevertheless, these items can sometimes be "individualized" and made into more discrete entities via listing. É. Kiss (1998: 262) provides the example presented here as (63), "where a two-member set of properties (including *tired* and *sick*) is established as a domain of exhaustive identification" (italic original). This technique allows the adjective *sick* to be accommodated into the focal position, creating a much more acceptable *it*-cleft than its counterpart in (62).

(62) *It's sick that he was

(63) It's not sick that he was but tired (É. Kiss 1998: 262)

By providing semantic rather than syntactic categorial constraints, É. Kiss' analysis correctly predicts that the range of *it*-cleft foci will not be comprised of complete phrasal categories.

[13] In Davidse's (2000) *it*-cleft analysis, the focal element is the antecedent to the relative clause. This may be one reason for Davidse's claim that the postcopular position is a "strictly 'nominal' slot". However, in the analysis provided in this thesis, the restrictive relative clause modifies the constituent *it*. As a result, non-nominal focal elements are not a problem to this analysis and we do not need to invoke type-shifting operations such as nominalization to accommodate them.

However, given this set of assumptions, we cannot yet explain why *it*-clefts have a different range of foci as compared with other specificational copular constructions.[14] As I go on to show in §8.2, the *wh*-cleft construction is in many ways more relaxed than the *it*-cleft about what can be accommodated into the referential slot, allowing a greater variety of predicative, property-denoting foci (including adjective phrases and verb phrases) to enter in the postcopular position. An example of this can be seen here by comparing the ungrammatical AP-focus *it*-cleft in (62) with the perfectly acceptable AP-focus *wh*-cleft in (61). I suggest that one way to account for these differences is to examine how these constructions have developed over time.

In §8.2, I show that although *it*-clefts and *wh*-clefts inherit from the same basic specificational schema, they have nonetheless undergone separate historical developments, acquiring construction-specific properties and functions over time. This discussion is based in part on the diachronic investigation of *it*-clefts outlined in §7.2. Here I show that *it*-clefts were at one time restricted to NP foci. As the *it*-cleft gradually accommodates non-nominal elements into the postcopular position via a series of incremental stages, the constructional schema becomes more abstract; that is, it loses a syntactic constraint. Instead, all that remains is a referentiality requirement; the perimeters of which are specific to the *it*-cleft construction and are dependent upon the *it*-cleft's peculiar discourse-functional properties.

5.4 'Informative-presupposition' (IP) *it*-clefts

The final type of (specificational) *it*-cleft that I discuss in this chapter is the 'informative-presupposition' *it*-cleft. In such examples, the relative clause expresses discourse-new, or unfamiliar, information. Since the cleft clause is usually associated with expressing discourse-old information and is sometimes said to contain a "familiarity presupposition", IP *it*-clefts are treated by many authors as atypical or marked examples. In what follows, I explain that there is good reason for this. While *it*-

[14] For É. Kiss (1998) however, *it*-clefts are not expected to exhibit the same range of foci as other specificational copular sentences. É. Kiss claims that "identificational focus" is dependent upon a particular syntactic configuration which, in English, is specific to the *it*-cleft construction. I strongly disagree with this aspect of É. Kiss' analysis, as shown in §4.1.2.3.

clefts containing discourse-old information are shown to inherit this property from other constructions, IP *it*-clefts require an independent explanation.

The association of the cleft clause with discourse-old information is predictable if we assume that *it*-clefts are a type of specificational copular sentence. Over the previous two chapters, I have made two important claims. First, I have argued that specificational meaning results from a special type of nominal predication relation involving definite noun phrase predicates. And secondly, I have advocated an analysis of *it*-clefts in which the constituent *it* and the relative clause function together as a discontinuous definite description. From this, it follows that *it*-clefts will inherit properties from the wider specificational copular construction and, consequently, from definite noun phrases in general.

It is well-known that definite noun phrases, like the relative clauses in *it*-clefts, are associated with familiar information. In §4.1.2.5, I claimed that the reason for this is that definite noun phrases involve existential commitment. It is not such a large step to assume that if the speech participants can assume that the set described by the definite noun phrase exists, then they are likely to be familiar with that set.

The association of the cleft clause with discourse-old information can therefore be traced back to the inherent semantics of definite noun phrases. But how can we explain the behaviour of *it*-clefts such as (64), in which the relative clause expresses discourse-new information that is not already known to the intended audience? Does this type of example, labelled the 'informative presupposition' *it*-cleft by Prince (1978), also inherit its behaviour from definite noun phrases?

(64) (Start of lecture)
It was Cicero who once said, 'Laws are silent at times of war'

For Hedberg (2000) and Gundel, Hedberg and Zacharski (2001), the existence of the IP *it*-cleft falls out straightforwardly from the behaviour of definite descriptions. Like me, these authors claim that the constituent *it* and the cleft clause in *it*-clefts function together as a definite-like expression. However, they argue that while definite

noun phrases are often claimed to be associated with familiar information, this generalization is not true in a large number of cases. For example, Gundel, Hedberg and Zacharski (2001) find that 44% of the definite article phrases in their corpus are in fact classifiable as 'nonfamiliar'. They conclude that the condition on the use of definite article phrases is that the referent described must be 'uniquely identifiable' rather than 'familiar'. This means that, at the very least, the addressee has to "be able to assign a unique representation to the phrase as soon as he or she has finished processing it" (Gundel, Hedberg and Zacharski 2001: 274).

Gundel, Hedberg and Zacharski (2001) claim that a parallel account can be provided for cleft sentences. Citing the corpus-based studies of Delin (1989) and Hedberg (1990), they note that the frequency of *it*-clefts containing new or inferable information in the cleft clause is relatively high, ranging between 49 and 72%. They conclude that, like definite noun phrases, the cleft clause is always uniquely identifiable, but does not have to be familiar. For these authors then, so-called IP *it*-clefts are not marked or exceptional uses. The fact that the cleft clause in *it*-clefts exhibits variation between discourse-old and discourse-new information is therefore expected and can be traced back to the behaviour of definite noun phrases in general.

There some important benefits to the analysis of IP *it*-clefts put forward by Hedberg (2000) and developed by Gundel, Hedberg and Zacharski (2001). For one thing, it allows a truly unified analysis of *it*-clefts with familiar and nonfamiliar information in the relative clause. Furthermore, this account is maximally explanatory; if we take a constructional perspective, we can say that both types of *it*-cleft inherit their behaviour from definite article constructions. As a result, IP *it*-clefts do not require an independent explanation.

Nevertheless, despite its appeal, there are methodological problems with this approach which, to my mind, cannot be overlooked. First, Hedberg (2000) and Gundel, Hedberg and Zacharski (2001) investigate the information status of definite noun phrases in a wide variety of contexts and do not focus on those occurring in specificational copular sentences. These authors are correct that when definite noun phrases are used referentially, the hearer may not always be familiar with that referent

(see also Donnellan 1966). For example, (65) can be uttered without the addressee having had any contact with, or explicit knowledge of, *the dog next door*. However, the definite-like expression in specificational *it*-clefts is not used referentially. Instead, it functions as the definite NP predicate of a specificational *NP be NP* sentence.

(65) The dog next door kept me awake (Hedberg 2000: 895)

Therefore, in order to ascertain whether IP *it*-clefts inherit the property of expressing hearer-new information in the cleft clause, we need to examine them in relation to corresponding specificational copular sentences containing definite descriptions. While the IP *it*-cleft in (64) above is a perfectly acceptable way to introduce a discourse, the corresponding *NP be NP* sentence in (66) seems strange and suggests that we should already know that *someone once said 'laws are silent at times of war'*. This suggests that while *it*-clefts can occur with hearer-new information in the relative clause, this property is not inherited from the wider specificational copular construction and is not predictable from the behaviour of definite noun phrases.

(66) (Start of lecture)
#The one/man who once said, 'Laws are silent in times of war' was Cicero

A further problem with the conclusions of Hedberg (2000) and Gundel, Hedberg and Zacharski (2001) is that their concept of 'nonfamiliar' encompasses both new and inferable information. For instance, (67) is cited as containing a definite NP which has the status of 'uniquely identifiable' rather than 'familiar'. Nevertheless, the existence of the referent *the bride* is inferable from the familiar script or frame of *wedding*. That is, we know from experiential or cultural knowledge that weddings include a bride (see also Clark and Haviland (1977) on 'bridging inferences').[15]

[15] Likewise, the existence of *the dog next door* in (65) is consistent with our cultural knowledge that dogs are common household pets.

(67) I've just been to a wedding. **The bride** wore blue.

(Lyons 1999: 7; as cited in GHZ 2001: 280)

It is not always clear where the boundary lies between given and new information. However, I suggest that by conflating the concepts 'new' and 'inferable', these authors fail to notice an important difference in the behaviour of *it*-clefts and other copular constructions. While all types of specificational copular sentence can occur with inferable information in the definite NP predicate, only *it*-clefts allow brand-new information to be expressed by the relative clause. For example, in (68), the specificational *NP be NP* sentence is introduced by a definite NP subject containing inferable information. Likewise, the corpus data of Prince (1978) and Collins (1991a) suggests that *wh*-clefts with inferable information in the relative clause are particularly common (see also §8.2).[16]

(68) I've just been to a wedding. **The bride** was Susan Aldridge.

In contrast, Prince (1978) and Collins (1991a) find that *wh*-clefts do not occur with brand-new information in the relative clause. This, coupled with the unacceptability of (66), suggests that IP *it*-clefts such as (64) require an independent explanation.

The distinction between 'new' and 'inferable' therefore seems to be particularly important for distinguishing between different types of specificational copular sentence. This calls into question Gundel, Hedberg and Zacharski's (2001) use of the data provided by Delin (1989) and Hedberg (1990). Since these frequency counts conflate *it*-clefts with both new and inferable information, they cannot be used as evidence that informative-presupposition *it*-clefts (defined by Prince (1978) as containing brand-new, rather than inferable, information) are unmarked or especially common.

As a result, the conclusions of Hedberg (2000) and Gundel, Hedberg and Zacharski (2001) are not supported by this reexamination of the data. The IP *it*-cleft does

[16] Other corpus studies of present-day English cleft sentences include Breivik (1986), Erdmann (1986), Geluykens (1988), Gómez-González (2004), Hasselgård (2004), Hedberg and Fadden (2007), Johansson (2001) and Nelson (1997).

not inherit its information-structure from the wider specificational copular construction and does not conform to the behaviour of definite noun phrases in general. Consequently, this type of *it*-cleft requires an independent explanation. In §7.2, I show that the IP *it*-cleft has emerged via extension from *it*-clefts with discourse-old information in the cleft clause. This involves the accommodation (or coercion) of new information into an inherently presuppositional constructional schema.[17] My diachronic story therefore explains where this construction-specific property comes from as well as offering a unified account of all types of *it*-cleft as emerging from a single source. The historical evidence therefore argues against accounts which claim that IP *it*-clefts make up a separate construction from the *it*-cleft proper.[18]

5.5 Chapter summary

Throughout this thesis, I have claimed that *it*-clefts are foremost specificational sentences. I argued that by maximizing inheritance from the wider specificational copular construction, we obtain the most explanatory analysis of *it*-clefts. However, in this chapter, I have highlighted three different types of *it*-cleft which exhibit properties that cannot be motivated by inheritance from this more basic construction: predicational *it*-clefts, non-NP focus *it*-clefts and informative-presupposition *it*-clefts.

I have shown that predicational *it*-clefts can be accommodated fairly neatly into my account. Rather than inheriting their semantics from the specificational copular construction, predicational (and proverbial) *it*-clefts inherit from the predicate nominal construction. As I explained in §3.3, specificational meaning is really just a special type

---

[17] This contrasts with the main claim of Gundel, Hedberg and Zacharski (2001) which states that since 'nonfamiliar' information is so common in *it*-clefts and definite NPs in general, such examples are not exceptional and do not result from a process of 'accommodation', as defined by Lewis (1979) and Heim (1982).

[18] For example, Ball (1991, 1994a) claims that the IP *it*-cleft developed out of a merger between several distinct constructions rather than via extension from existing *it*-cleft constructs. In §7.2, I show that this account is not supported by the historical evidence. Likewise, Gundel (1985) claims that since *it*-clefts are associated with given information, IP *it*-clefts form a separate construction with a different thematic structure. In §8.2, I show that it is precisely this association between *it*-clefts and discourse-old, rather than inferable, information that lends itself to pragmatic accommodation. Finally, Declerck (1984a) argues that IP *it*-clefts such as (64) above have an emphasizing rather than specifying function. However, in §8.2, I show that while these examples have additional discourse functions, they nevertheless retain a specificational meaning.

of nominal predication relation. From this, it follows that *it*-clefts with nominal foci can be given two different interpretations. If the discontinuous definite description is analysed as the semantic predicate, a specificational reading ensues. However, if the focal NP is analysed as the semantic predicate and the definite-like element is used as a referring expression, then the result is a predicational *it*-cleft.

The specificational-predicational distinction in *it*-clefts is therefore inherited from other copular constructions. However, I have shown that the range of *it*-cleft foci and the information status of the cleft clause are construction-specific properties which are not inherited from the language system. In §7.2, I undertake a diachronic investigation into the development of non-NP focus *it*-clefts and IP *it*-clefts. Here, I show that historical evidence can be used to provide motivation, and an explanation, for their idiosyncratic properties.

6. THE ADVANTAGES TO A CONSTRUCTIONAL ACCOUNT OF *IT*-CLEFTS

Over the last three chapters, I have outlined a synchronic analysis of *it*-clefts and have shown how this compares favourably to the alternative accounts proposed in the literature. In this chapter, I reflect on how the theoretical framework of construction grammar has played an instrumental role in shaping my analysis. In §6.1, I reintroduce some of the principles and concepts that are specific to a constructional, usage-based theory of grammar and show how I have made use of them. In §6.2, I go on to compare my analysis to other so-called "constructional" *it*-cleft accounts proposed in the literature. As I explain, the analysis put forward in this thesis is more in keeping with the basic claims that underlie the constructional framework and makes better use of the machinery that provides a constructional approach with explanatory adequacy. I conclude that the analysis presented in this thesis is therefore more "constructional" than alternative *it*-cleft analyses.

In §6.3, I ask what the advantages are to conducting an analysis of *it*-clefts within the framework of construction grammar. I argue that a constructional approach is better suited to describing and explaining the *it*-cleft's properties than other frameworks, including both generative and transformational theories of grammar. Finally, §6.4 provides a constructional hierarchy of cleft and specificational sentences based upon the familial relationships identified in previous chapters. This serves as a summary of our progress so far and highlights the areas that are outside the scope of inheritance. These as yet unexplained properties provide the basis of the diachronic study undertaken in chapters 7 and 8.

6.1 What is constructional about my account?

As I explained in chapter 2, construction grammar was developed with a mind to providing full and explanatory analyses of specialized linguistic patterns, which cannot be accounted for by appealing to highly general rules. In order to achieve this goal, each linguistic pattern is realized as a separate node within a taxonomic network of constructions. By viewing the grammar as a network, rather than as a set of fixed and

finite rules, constructional theories offer a unique way of viewing the relationship between the different units of grammatical knowledge. This network of constructions has a hierarchical structure, showing that some linguistic patterns are more basic or general than others. Much of the explanatory power of construction grammar comes from the premise that lower level constructions inherit attributes from higher-level, or more schematic, constructions. The theory is that the more properties that are inherited from the language system, the more "motivated" that construction is. In other words, if a construction fits well into the overall linguistic system, we have an explanation as to why this particular construction, with these particular form-meaning correspondences, should exist.

The analysis built up throughout chapters 3, 4 and 5 aims to maximize motivation for the *it*-cleft construction. Starting from the premise that *it*-clefts are specificational copular sentences, I provide an analysis that makes important generalizations in an inheritance hierarchy. In §4.1, I advocated an extraposition-from-NP analysis of *it*-clefts in which the initial *it* and the sentence-final clause operate together like a discontinuous definite description. This allows me to analyse all specificational copular sentences, including the *it*-cleft, as involving a nominal predication relation and explains how the shared pragmatic characteristics of presupposition, exhaustiveness, contrast, and discourse familiarity result ultimately from the inherent semantics of definite noun phrases. In this analysis, the *it*-cleft is therefore firmly situated within a family of related constructions and the maximum number of its properties are explained as being inherited from the more general specificational copular schema.

This analysis therefore shows that while the *it*-cleft construction is noticeably marked with respect to highly general rules of grammar, many of its attributes can be explained if we examine how *it*-clefts relate to structurally similar constructions in the network. However, construction grammar recognizes that not all of the properties of a given construction are inherited from more schematic linguistic patterns. This represents a major distinction between construction grammar and the movement-based theories of the generative tradition. In movement-based theories, the assumption is that surface

idiosyncrasies (or mismatches) are derived from underlying structures which behave in expected ways and conform to more general rules of grammar (see §2.4). In contrast, construction grammar, which does not rely on a hard and fast system of rules, permits and expects construction-specific properties. An important advantage to this philosophy is that construction grammar does not need to explain away idiosyncrasies via complex and counterintuitive derivations; psychologically, the constructionist concern is for reducing demands on processing rather than on creating a maximally economical system (see Goldberg 1995: 74; Croft and Cruse 2004: 278).

The architecture of constructions as signs, that is, as form-meaning pairs, makes construction grammar particularly adept at incorporating idiosyncratic information into the grammar. In this model, the symbolic relations which link units of form and meaning are internal to the construction, rather than imposed by general rules. As a result, information in the construction may 'override' inheritance from more general patterns of correspondence, creating 'mismatch'. I make use of this property in the analysis of specificational sentences given in §3.3.3, where I argue that they involve a nominal predication relation. In canonical specificational sentences, such as (1), the semantic predicate precedes the referring expression.

(1) The murderer is John

In this construction then, we have a mismatch between the syntax and the semantics. The generalization that the syntactic subject and its complement correspond to the semantic argument and predicate, respectively, is pervasive. However, in the canonical specificational copular construction, the syntactic subject is also the semantic predicate. In contrast to the movement-based inverse analysis outlined in §3.2.2, a monostratal constructional account does not provide such examples with an underlying structure that conforms to the subject-predicate pattern. Instead, the specificational construction simply overrides this generalization.

By accepting idiosyncrasies in the grammatical system, construction grammar does not lose any of its explanatory force. In fact, as shown by the analysis of *it*-clefts

put forward in §4.1, it is often what allows a constructional account to make the relevant generalizations. For example, my analysis assumes that the *it*-cleft contains a non-standard use of the pronoun *it*, which functions both as the determiner and nominal head of a definite description. Likewise, by assuming that the sentence-final clause restrictively modifies the constituent *it*, I overlook the strong tendency for restrictive relative clauses to modify an immediate antecedent and the fact that the pronoun *it* cannot be restrictively modified outside the context of the *it*-cleft construction. However, providing these properties with a construction-specific analysis enables my account to recognize the similarities between *it*-clefts and not only other specificational copular sentences, but also definite noun phrases and restrictive relative clauses as well. As I have shown, the explanatory significance of the inheritance relations I identify far outweighs the untidiness of acknowledging a handful of idiosyncrasies in the *it*-cleft construction.

Furthermore, while idiosyncrasies are tolerated in a constructional model of language structure, every effort is made to provide them with a proper description and explanation. Goldberg (2003: 120) notes that "What imbues a constructional approach with explanatory adequacy is a further desideratum that each construction must be *motivated*" (italics original). While construction-specific properties cannot be motivated by inheritance, motivation can be provided by factors external to the grammar. In chapters 7 and 8, I go on to show that historical evidence can help to provide motivation for and explain the construction-specific properties of the *it*-cleft. As I explain below in §6.3, one of the advantages of usage-based constructional theories, is that they are able to intersect with and work alongside theories of language change.

6.2 What makes a good constructional account?

In the previous section, I have shown that my analysis of *it*-clefts and specificational copular sentences conforms to the principles of construction grammar and makes full use of the concepts and mechanisms that make it up. In this section, I compare my account to other *it*-cleft analyses that claim to be "constructional" in some sense, namely Lambrecht (2001), Davidse (2000) and Pavey (2004). I show that while each of these

accounts makes reference to some of the basic claims that underlie the constructional framework, none of them is successful at meeting the explanatory objective that for each construction, motivation must be maximized (Goldberg 2003: 120). I conclude that the analysis of *it*-clefts sketched in this thesis is more compatible with the fundamental principles of construction grammar.

6.2.1 Lambrecht (2001)

Lambrecht (2001) proposes a constructional account of *it*-clefts that is largely based on Jespersen (1937). To recap from §3.1, Jespersen (1937) provides an analysis which emphasizes the relationship between *it*-clefts and their noncopular counterparts. For example, according to Jespersen's (1937) analysis, the *it*-cleft in (2) has the same underlying predication relationship as the subject-predicate sentence in (3).

(2) [It was] Howard [that] left

(3) Howard left

For Jespersen then, the only difference between these two examples lies in their information structure. In (2), the informational content is "cleaved" or split up by the addition of the semantically empty elements *it*, *be* and *that*, with the result that the referential subject is placed in a syntactically marked focal position.

In keeping with construction grammar's monostratal architecture, Lambrecht dismisses the derivational aspects of Jespersen's analysis, suggesting that the postcopular element is not an underlying subject. However, Lambrecht (2001: 472) agrees with Jespersen's claim that the sequence *it is* and the connective marker "do not enter fully into the semantic composition of the sentence". For Lambrecht then, despite its biclausal syntax, the *it*-cleft expresses a simple subject-predicate proposition. He suggests that the semantic subject is expressed by the gap in the sentence-final clause, which is "coindexed with the predicative argument of the copular" (Lambrecht 2001: 467), as shown in (4).

(4) [[It was [Howard]$_i$] [that ____$_i$ left]]

In Lambrecht's analysis, the postcopular focal element is therefore the antecedent to the cleft clause. Since the sentence-final clause in *it*-clefts does not modify its antecedent, Lambrecht classifies it as a special type of nonrestrictive relative. As I explained in §4.1.3, the cleft clause exhibits a number of properties that pose a problem for a nonrestrictive relative clause analysis. However, rather than claiming that the cleft clause patterns exactly like other nonrestrictive relatives, Lambrecht (2001) simply argues for a unitary analysis of relative clauses which can encompass those found in cleft sentences.[1] He notes, "I take as the fundamental property of all RCs that they are PREDICATES" (Lambrecht 2001: 473; emphasis original). Under this extremely broad definition, the cleft clause can be classified as an, albeit non-modifying, relative, which is predicated of the postcopular semantic subject.

Lambrecht's analysis is therefore a non-derivational and monostratal variant of the expletive approach. As I explained in §3.1, such analyses prioritize the relationship between the postcopular focal element and the sentence-final clause, with most authors assuming that the cleft clause is directly predicated of the clefted constituent. Since, in this approach, the constituents *it* and *be* do not enter into the semantics of the utterance, expletive accounts emphasize the correspondence between *it*-clefts and noncopular subject-predicate sentences. As a result of this, Lambrecht (2001: 472) assumes that the *it*-cleft's main function is as a syntactic focusing device, with *it* and *be* providing the pragmatic role of "focus marker". In particular, Lambrecht (1994, 2001) suggests that clefts provide an alternative way of marking 'argument-focus'. For example, (5) involves the prosodic focus marking of the argument *Howard*, whereas the cleft in (6) places the argument in a syntactically marked focus position.

(5) HOWARD left

(6) It was Howard that left

[1] Lambrecht's (2001: 273) "unitary" analysis is formulated in response to (what he believes is) the widespread assumption that relative clauses are "restrictive by definition".

Following Declerck (1988), Lambrecht (2001: 496) assumes that argument-focus is the same thing as specificational meaning. For example, using Declerck's (1988) 'value-variable' terminology, both (5) and (6) can be said to provide the value *Howard* for the variable in *x left* (see also Declerck 1988 and the discussion of his work in §3.2.3).

Lambrecht claims that his analysis conforms to the basic principles of construction grammar. He notes that the big advantage of this particular grammatical theory is that it allows for noncompositional linguistic patterns (Lambrecht 2001: 466). Consequently, construction grammar does not in principle rule out the highly idiosyncratic analysis of *it*-clefts that Lambrecht proposes. For example, although Lambrecht does not use this term, his analysis involves a complexity mismatch. Francis and Michaelis (2003: 5) define a 'complexity mismatch' as occurring when there is no one-to-one correspondence between the formal elements and the semantic components of a construction (see §2.4). For Lambrecht (2001: 466), the construction is syntactically complex but semantically simple, involving a number of dummy elements; he claims that cleft sentences "express a simple proposition via biclausal syntax".

Likewise, although Lambrecht does not discuss this issue in detail, his analysis presupposes a clausal constituent that is unique to cleft sentences. As I explained in §4.1.3, there are different criteria which characterize restrictive and nonrestrictive relative clauses. If the antecedent to the cleft clause is understood to be the complement of *be*, as it is in Lambrecht's (2001) expletive account, then the sentence-final clause in *it*-clefts patterns with neither of these relative clause types. Therefore, while Lambrecht (2001: 472) claims that the cleft clause is a type of relative, it must nevertheless be distinct from both restrictive and nonrestrictive relative clauses occurring outside of the *it*-cleft construction. Consequently, although Lambrecht (2001) does not make this claim directly, his analysis requires the stipulation of a type of clause that, in terms of its relationship to its antecedent, is unique to the *it*-cleft construction.

In addition, although Lambrecht acknowledges that *it*-clefts contain existential and exhaustiveness presuppositions, he does not explain either how these properties come about or which of the *it*-cleft's components contribute to them. If these meanings cannot be attributed to any of the individual components, we must assume that they are

associated with the construction as a whole. The existence of noncompositional meaning of this sort does not pose a problem for construction grammar. As I explained in §2.5, in this grammatical theory, symbolic relations, which link units of form and meaning, are internal to the construction. As well as linking say, the form of a single lexical item to a single component of meaning, symbolic relations also link the construction's complete structure to aspects of the construction's conventional meaning. This means that a construction can have meaning that is associated with the construction as a whole but which cannot be attributed to any one of its individual elements.

Lambrecht's analysis is therefore constructional in the sense that he believes, as I do, that the concept and architecture of the grammatical construction is extremely useful for configurations, like the *it*-cleft, which exhibit properties that are not predictable from highly general rules of the grammar. However, despite this, Lambrecht's *it*-cleft analysis is hugely different from my own constructional account. The reason for this is that while Lambrecht treats the *it*-cleft as a construction, he does not try to ensure that this construction is fully "motivated" and situated in relation to the rest of the language system. For example, Lambrecht (2001: 466) claims that constructions, such as the *it*-cleft, which contain properties that are not predictable from universal grammatical rules, "require independent explanation". However, by studying a linguistic pattern in isolation and providing a highly idiosyncratic analysis, little or no explanation is provided as to why this particular configuration should exist in a language that otherwise seems not to support it.

In his 1994 book, Lambrecht does provide some explanation for the *it*-cleft configuration in the form of general principles of information-structure. For example, he notes that there is a strong tendency for the subject of a sentence to express given, or discourse-old, information. Lambrecht (1994: 22) comments, "The pragmatic function of the clefting structure is to create an additional postverbal argument position in which the focus NP may appear, preventing it from occurring in sentence initial position". However, information-structure principles cannot alone explain why this particular construction, with these particular form-meaning correspondences, is used to mark focus. For example, it isn't clear from this analysis why the constituents *it*, *be* and *that*

should be present in the syntax if they do not contribute to the construction's meaning. Why should the elements *it* and *be* in particular be used as focus markers? How did the construction develop exhaustiveness and existential presuppositions? Why does this construction exhibit a unique type of relative clause?

Lambrecht's (2001) analysis therefore leaves so many questions unanswered and stipulates so many construction-specific properties that it becomes a largely descriptive, rather than explanatory, account. While Lambrecht makes use of the concept and architecture of the grammatical construction, he does not discuss the organization of the language system as a hierarchical network. As I go on to show in §6.2.4, the expletive account advocated by Lambrecht is not supported by the way that constructional taxonomies are organized. Therefore, I conclude that Lambrecht's (2001) analysis is only loosely "constructional" and does not achieve the main goal of construction grammar: to provide full and explanatory accounts of specialized linguistic patterns.

6.2.2 Davidse (2000)

Davidse's (2000) constructional account differs considerably from Lambrecht's (2001) analysis. Unlike Lambrecht, Davidse does not assume that *it*-clefts are semantically equivalent to their noncopular subject-predicate counterparts. Instead, she claims that both the matrix clause and the cleft clause contribute to the *it*-cleft's representational semantics. For Davidse, the matrix clause is an "identifying clause"; in other words, it functions as a specificational copular sentence. In Davidse's analysis then, the matrix clause of the *it*-cleft in (7) has the same analysis as the *NP be NP* specificational sentence in (8)[2].

(7) [It was Howard] that left

(8) It was Howard

[2] Davidse (2000) assumes, as I do, that example (8) is an *NP be NP* specificational sentence introduced by the semantically meaningful pronoun *it*. However, for authors who analyse *it*-clefts as containing an expletive pronoun, such sentences are often analysed as "truncated clefts" (see §4.1). From this perspective, example (8) is an *it*-cleft whose relative clause is "recoverable" from the discourse context.

Using Halliday's (1967) terminology, Davidse (2000) labels the pronoun *it* as the 'identified' and the postcopular element (in this case *Howard*) as the 'identifier'. In many ways, the 'identifier-identified' relationship is akin to Declerck's (1988) 'value-variable' relation (see §3.2.3). As Davidse (2000: 1121) comments, "The identified can be likened to the unknown "x" in a mathematical equation and the identifier to its actual value in that equation".

In addition, Davidse (2000) claims that the constituent *it* performs a quantifying role in the matrix clause. She argues that since *it* is an inherently definite pronoun, it "conveys quantificational exhaustiveness" (Davidse 2000: 1121). By equating the universally quantifying *it* with the postcopular constituent, the matrix clause imposes an exhaustively specifying value onto the complement of *be* (Davidse 2000: 1125). In this way, Davidse accounts for the exhaustiveness presupposition of *it*-clefts. The relative clause then takes this complement (which has been exhaustively quantified) as its antecedent, incorporating it into a value-variable relation, as defined by Declerck (1988). In other words, the postcopular referring expression constitutes the specific 'value', or total set of instances, for the 'variable' or "entity involved in the situation designated by the relative clause" (Davidse 2000: 1125).

For Davidse, the cleft clause is neither a restrictive nor nonrestrictive relative clause. She notes that while restrictive relatives modify and restrict the "type specification" of the head noun of a noun phrase, the cleft clause incorporates a full NP antecedent into a value-variable relationship. On the other hand, like restrictive relatives, the cleft clause can be introduced by *that* or zero, elements which cannot introduce a nonrestrictive relative clause (see §4.1.3). Furthermore, as I noted in §5.5, Davidse (2000: 1116) claims that, while nonrestrictive relatives can have non-nominal antecedents, the antecedent to the cleft clause is always an NP "either as such or as the result of the reclassification of non-nominal units". That is, while *it*-clefts allow a range of elements to occur in the postcopular position, Davidse (2000: 1116) argues that they are nominalized into what is essentially a "strictly 'nominal' slot". Davidse (2000: 1128) therefore concludes that the relationship between the cleft clause and its antecedent has a "special status".

Like Lambrecht (2001) then, Davidse (2000) assumes that the antecedent to the cleft clause is the complement of *be*. However, rather than claiming that *it*-clefts express the same semantic content as their noncopular counterparts, she relates *it*-clefts, as I do, to other specificational copular sentences. As a result, Davidse provides the *it*-cleft with a much more compositional structure than Lambrecht (2001). For instance, while Lambrecht's analysis involves a complexity mismatch between the syntax and the semantics of the *it*-cleft construction, Davidse assigns a semantic function to each of its syntactic elements. In an analysis similar to my own, Davidse (2000: 1121) argues that the pronoun *it* is not a semantically empty element and instead performs an important quantificational role. She notes, as I do, that the initial *it* in *it*-clefts behaves in a similar way to the definite article in *th*-clefts; both *it* and *the* involve proportional (or universal) quantification, providing the cleft sentence with its exhaustiveness presuppositions (see §4.1.2.3). This contrasts with Lambrecht (2001), who assumes that the cleft pronoun is semantically empty and does not provide any explanation as to where the exhaustiveness of *it*-clefts comes from.

On the surface then, Davidse's account appears to be more explanatory than Lambrecht's. However, in many ways, her analysis is just as idiosyncratic and also suffers from a number of inconsistencies. For example, although Davidse views the matrix clause of *it*-clefts in relation to other "identifying clauses" (or specificational *NP be NP* sentences) she nevertheless comes up with an analysis of the relative clause that is unique to cleft sentences. If this type of clause (which incorporates its antecedent into a value-variable relation) is not inherited from more general constructions in the language system, then where does it come from and how is it motivated?

In addition, Davidse's analysis is unnecessarily complex. As I explained above, Davidse's account does not conform to the major principles associated with the expletive approach. However, it is also distinct from extraposition accounts, including the analysis presented in this thesis. She notes that while extraposition-from-NP analyses reduce *it*-clefts to monoclausal specificational sentences, her analysis assumes that *it*-clefts are biclausal, with each clause entering into a different semantic relationship with the postcopular element. Davidse (2000: 1127) suggests that this is advantageous, since

the function of *it*-clefts is not simply to express specification, but to specify "INSTANCES AS VALUES CORRESPONDING TO A VARIABLE" (emphasis original). For Davidse then, the exhaustive specification of the postcopular element is the function of the matrix clause, while the relative clause incorporates it into a value-variable relationship.

However, to my mind, it seems that Davidse chooses a complex analysis over a simple one without adding anything to the explanatory significance of her account. As I showed in §4.1, if we analyse *it*-clefts as specificational sentences containing discontinuous definite descriptions, we are able to account for the construction's identifying meaning and exhaustiveness presuppositions, among other properties. Since this simple analysis is maximally explanatory, why do we need the more complex account that Davidse proposes? On Davidse's account, specificational meaning is marked twice: once by the identifying matrix clause and once by the value-variable relationship expressed by the relative clause. Aside from being uneconomical, this is also unnecessary, since it increases the complexity of the *it*-cleft structure in a way that is not supported by inheritance from the language system.

Davidse's account also suffers from an important inconsistency. She assumes that the constituent *it* exhaustively quantifies over the postcopular element **before** it serves as the antecedent to the relative clause. However, if the quantifying *it* does not have scope over the relative clause, then how can quantification in the matrix clause lead to exhaustiveness presuppositions in the value-variable relationship of the relative clause? As Pavey (2004: 76) comments, "Davidse repeats several times that the ordering, or scope, of the two semantic relationships she posits is significant but she does not explain this ordering". Furthermore, while Davidse (2000: 1121) claims, as I do, that the *it* in *it*-clefts performs a similar quantificational role to the definite article, her analysis actually requires that the cleft pronoun and the definite determiner behave differently. For example, while the definite article has scope over the head noun and its modifiers, including relative clauses, the cleft pronoun in Davidse's analysis only operates on the postcopular element (Pavey 2004: 76). Consequently, despite Davidse's claims, her analysis of *it*-clefts is not supported (or motivated) by the behaviour of definite noun phrases in other identifying clauses.

6.2.3 Pavey (2004)

Unlike Lambrecht (2001) and Davidse (2000), Pavey does not conduct her analysis in the framework of construction grammar. Nevertheless, there are a number of reasons for including Pavey (2004) in a review of constructional *it*-cleft analyses. For one thing, Pavey's account is formulated within a Role and Reference Grammar framework. Like construction grammar, this is a non-derivational, monostratal, parallel architecture theory, which allows for the intersection of syntactic, semantic and pragmatic representations. In addition, and most importantly for our purposes, Role and Reference Grammar makes use of the concept of the construction as a unit of grammatical knowledge[3]. As a result, this theory permits and expects the grammar to contain idiosyncratic information. As Pavey (2004: 246) comments, construction-specific properties and conditions are stated in the constructional template.

The framework in which Pavey conducts her analysis therefore shares a number of similarities with construction grammar. In addition, and partly as a result of this, Pavey's analysis shows a strong resemblance to the constructional accounts of Lambrecht (2001) and Davidse (2000). For instance, Pavey (2004:76) notes that "Davidse's constructional approach makes useful advances and begins to bear similarities to the way a Role and Reference Grammar analysis would interpret the [*it*-cleft] construction". As a result, I include Pavey (2004) in my review of "constructional" *it*-cleft analyses. While it is unfair to criticize Pavey's account for not conforming to the basic claims and principles of construction grammar, my aim here is to show how her analysis differs from my own.

Pavey (2004) is heavily influenced by Declerck's (1988) understanding of *it*-clefts and other specificational sentences. As I explained in §3.2.3, Declerck argues that specificational meaning involves a value-variable relationship. For instance, in

[3] There are, of course, a number of important differences between Role and Reference Grammar and construction grammar. For instance, while in Role and Reference Grammar, the syntax is separate from the lexicon, construction grammar views all items of linguistic knowledge, including lexical items as constructions. In this way, all grammatical knowledge has a uniform representation and is stored in the same hierarchical system (see §2.3). As I go on to show in chapter 7, this aspect of the constructional framework means that it is well-suited to historical language study; it predicts that there is a continuum between the syntax and the lexicon, which is empirically supported by work in grammaticalization.

specificational *NP be NP* sentences, such as (9), the subject NP contains a variable *x* (as in *the x who committed the murder*) for which the postcopular NP (*John*) specifies the value.

(9) The one who committed the murder was John

Likewise, for Declerck, *it*-clefts too involve a value-variable relation. However, Declerck (1988: 185) notes that *it*-clefts differ from *NP be NP* specificational sentences because they do "not express the variable NP completely". For instance, in (9) above, the variable *x* is explicitly represented by the head noun of the subject NP (in this case, *one*). However, while Declerck claims that this same variable is presupposed to exist in the corresponding *it*-cleft in (10), the *x* variable is not represented in its syntactic form. In *it*-clefts then, the postcopular element provides the value for the variable **described by** the sentence-final clause.

(10) It was John who committed the murder

Pavey builds upon this analysis of *it*-clefts and presents it within a Role and Reference Grammar framework. Following Davidse (2000), Pavey claims that while the cleft clause has the same internal structure as a restrictive relative, it differs in its external relationship to the rest of the sentence. She notes that since the immediate antecedent to the cleft clause is a full noun phrase, this clause does not function as a restrictive modifier. Pavey represents this syntactically by "linking" the cleft clause to the matrix core (shown in Figure 6.1). Like Lambrecht (2001) and Davidse (2000) then, Pavey provides a construction-specific analysis of the cleft clause. She notes that "the relationship between the cleft clause and the rest of the *it*-cleft construction results in a "sens global particulier" [unique overall meaning] that will be characterized differently from both restrictive and non-restrictive relative clauses" (Pavey 2004: 201).

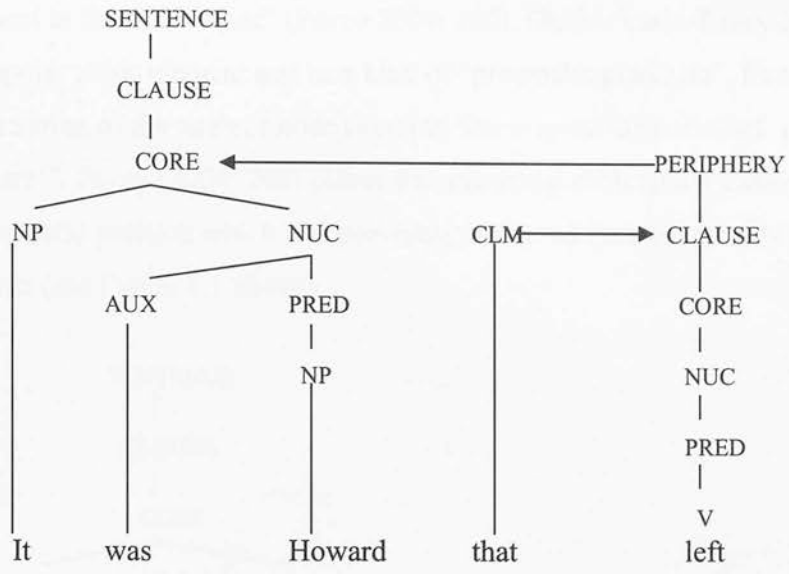


Figure 6.1 *Pavey's Role and Reference Grammar analysis (adapted from Pavey 2004: 207)*

For the matrix clause, Pavey adopts a simple predicational structure. In Role and Reference Grammar, each clause contains both a 'syntactic nucleus' (NUC) and a 'core argument'. The syntactic nucleus is reserved for the semantically predicative element. For example, in (11), the predicate nominal *a doctor* is the syntactic nucleus and its semantic argument *John* is placed in the 'core argument' slot. As shown in Figure 6.2, the copular verb *be* is labelled as an auxiliary.

(11) John is a doctor

However, this clause structure is not well-suited to the matrix clause in *it*-clefts, since the postcopular constituent is a semantically referring expression, rather than a predicative element. In order to get around this, Pavey claims that Declerck's 'value-variable' relation is a type of **pragmatic** predication relationship. Pavey observes that while in noun phrases, the function of a restrictive relative clause is to assist the hearer in identifying the referent of the head noun, the opposite relationship is found in *it*-clefts. She notes that the postcopular "value 'modifies' the variable expressed by the cleft clause [and]…assists the hearer in making a full identification of the underspecified

argument in the cleft clause" (Pavey 2004: 200). On this basis, Pavey claims that the postcopular value element acts as a kind of "pragmatic predicate". Extending the interpretation of the nucleus node (NUC) to "encompass the notion of 'pragmatic predicate'", Pavey (2004: 268) places the referential postcopular element of the *it*-cleft in a syntactic position which was previously reserved for semantically predicative elements (see Figure 6.1 above).

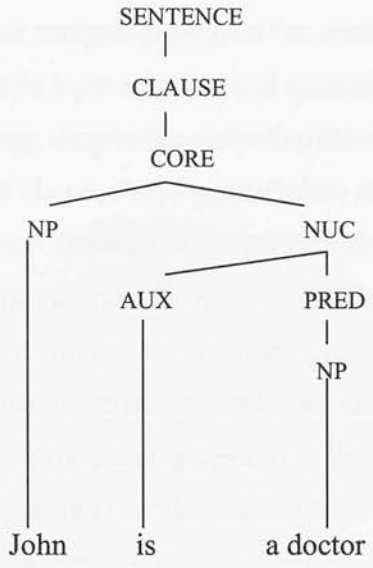


Figure 6.2 *Layered structure of the clause with a nominal predicate and* be *as auxiliary (adapted from Pavey 2004: 100)*

The cleft pronoun is placed in the 'core argument' slot, a syntactic position typically filled by the semantic argument of the predicative element (see Figure 6.2). Since on her analysis, the constituent *it* is not a semantically referring expression and is in fact altogether absent from the semantic representation of the construction, Pavey's analysis presents yet another mismatch between the syntax and the semantics. Nevertheless, Pavey argues that the pronoun *it* in *it*-clefts is not entirely meaningless. She claims that the cleft pronoun reflects the main **specificational** function of the

sentence – a relationship that is specified in the constructional template along with other idiosyncratic information[4].

Pavey's (2004) analysis of the *it*-cleft construction is therefore highly idiosyncratic. For instance, like Lambrecht (2001) and Davidse (2000), she stipulates a special type of syntactic and semantic relationship between the cleft clause and its antecedent which is not found outside of the *it*-cleft construction. In addition, Pavey's analysis is marked by the fact that it contains massive amounts of mismatch. She notes that her analysis highlights "an overriding mismatch, or lack of iconicity, between the semantic representation and syntactic form of the construction" (Pavey 2004: 289). For instance, despite her claim that there is no semantic predication relationship in the matrix clause, Pavey nevertheless provides it with a syntactically predicational clause structure. In order to incorporate the postcopular referring expression into a syntactic position reserved for predicative elements, Pavey provides an *ad hoc* story in which she labels it "pragmatic predicate". As Pavey (2004: 254) comments, "The matrix clause contains no verbal predicate; the nucleus contains a referring expression and the only argument (the cleft pronoun) is non-referring…These idiosyncrasies of the construction require special conditions and these are stated in the constructional template".

Pavey's (2004) reliance on mismatch and construction-specific phenomena is somewhat unexpected since her account originates from the same starting point as my own analysis: with the premise that *it*-clefts are a subtype of specificational copular

---

[4] This aspect of Pavey's analysis stems from her discussion of *there*-clefts. Pavey (2004: 155) claims that *there*-clefts serve to **introduce** the postcopular referent rather than **identify** it as the value for a variable. She concludes that while *it*-clefts are specificational, *there*-clefts are existential and argues that the purpose of the cleft pronoun is therefore to reflect the main semantic/pragmatic function of the sentence (Pavey 2004: 157). However, Pavey's analysis of *there*-clefts is based upon the wrong set of data and includes the existential sentence given in (i). Here, *a tall man who constantly falls over* is asserted to exist in the yoga class. Therefore, in this example, the relative clause restrictively modifies the antecedent noun. Such sentences should not be conflated with *there*-clefts. As Kearns (2000: 81-85) comments, since definite NPs are already presupposed to exist, existential sentences only allow indefinite NPs into the postcopular position. Consequently, these examples represent a separate construction from *there*-clefts such as (ii). As I explained in §4.1.2, *there*-clefts are specificational; they differ from *it*-clefts only in the fact that they are not exhaustive.

(i) (All kinds of new people have signed up for my yoga class…)
There's a tall man who constantly falls over… (Pavey 2004: 170)

(ii) Well, there's John and Mary that are available to work Saturday…

sentence. For instance, Pavey (2004: 184) notes that the *it*-cleft should be viewed "as one of a 'family' of copular constructions, and as a specificational construction in particular". However, while I have argued that specificational copular sentences involve a special type of nominal predication relation, Pavey assumes Declerck's (1988) account of specificational meaning as involving a value-variable relationship. As I explained in §4.1, my analysis of specificational meaning leads me to argue for an extraposition-from-NP analysis of *it*-clefts. In contrast, Declerck (1988) argues that while *it*-clefts share the same value-variable relationship as their corresponding noncleft specificational sentences, they are formally distinct. For instance, as I noted above, Declerck claims that the variable *x* in *it*-clefts is not syntactically realized. Pavey (2004: 187) concludes that all "specificational constructions share a 'pragmatically predicational' function while the *it*-cleft construction is distinguished by its syntactic form".

The analyses of Declerck (1988) and Pavey (2004) therefore fail to integrate *it*-clefts into a truly unified account of specificational copular sentences. Rather than explaining how the *it*-cleft's formal properties are motivated, Pavey's analysis highlights the idiosyncratic and noncompositional nature of *it*-clefts. Only the functional correspondences between *it*-clefts and other specificational constructions are emphasized. As I explain in §6.2.4, "families" of constructions are based upon formal resemblances. Therefore, Pavey's (2004) "constructional" account is not supported by the organization of the grammar in constructional theories.

6.2.4 A good constructional account?

Over the previous three subsections, I have discussed the *it*-cleft analyses of Lambrecht (2001), Davidse (2000) and Pavey (2004). I have shown that each of these authors makes reference to the basic claims and concepts that underlie the construction grammar framework. For instance, in these accounts, the *it*-cleft is acknowledged as having the status of a construction; that is, it is treated as a conventional unit of linguistic knowledge containing construction-internal mappings between form and meaning. In this sense then, the analyses of Lambrecht, Davidse and Pavey can be labelled "constructional". However, despite this, not one of these analyses shows a strong

similarity to the account of *it*-clefts argued for in this thesis. In what follows, I explain that the reason for this is that Lambrecht, Davidse and Pavey do not acknowledge or address the methodological objective for which construction grammar was designed: to provide full and explanatory accounts of both general and specialized linguistic patterns.

The analyses proposed by Lambrecht (2001), Davidse (2000) and to some extent Pavey (2004) make use of construction grammar primarily as a theory that tolerates idiosyncratic structures. For these authors, the term 'construction' is therefore synonymous with linguistic patterns consisting of form-meaning correspondences not found in other areas of the grammar and encoding constructional meanings that cannot be attributed to their component parts. However, although idiosyncrasies and exceptions are permitted in construction grammar, the explanatory power of the framework is based upon the organization of the language system as a hierarchical network of constructions. These constructions are related to one another via inheritance links, with specific constructions inheriting properties from more general patterns. The more properties a construction inherits, the more it can be said to be motivated by the language system.

Surprisingly, Lambrecht, Davidse and Pavey do not exploit inheritance in their analyses. Of course, this is not to say that these authors study *it*-clefts in complete isolation from other constructions[5]. For example, as shown in §6.2.1, Lambrecht's (2001) expletive analysis views *it*-clefts in relation to truth-conditionally equivalent subject-predicate sentences. However, Lambrecht's approach is not supported by the way that constructional taxonomies are organized. Goldberg (1995: 108) says that inheritance links are only posited between constructions that are formally related. Consequently, despite the similarity in meaning between clefts and their noncopular counterparts, this relationship is not represented in the grammatical system. This means then that the expletive approach is not supported by the framework of construction grammar. In contrast to Lambrecht, Pavey (2004) argues, as I do, that *it*-clefts are

[5] As a further example of this, Lambrecht, Davidse and Pavey add weight to their analyses by applying them to other types of cleft sentence, including *wh*-cleft and *th*-cleft examples (Lambrecht 2001) as well as clefts introduced by *there* (Davidse 2000 and Pavey 2004). However, since these constructions are superficially just as puzzling and peripheral as the *it*-cleft, such patterns of correspondence are not, on their own, helpful in providing motivation for the *it*-cleft construction.

members of the family of specificational copular constructions. However, as I noted in §6.2.3 above, Pavey only emphasizes functional, rather than formal, correspondences between *it*-clefts and other specificational sentences.

A truly constructional account of *it*-clefts seeks to maximize motivation for the *it*-cleft construction. For instance, my extraposition-from-NP analysis emphasizes both formal and functional correspondences between *it*-clefts and the specificational copular schema. Like noncleft specificational sentences, I analyse *it*-clefts as involving a nominal predication relation involving a (discontinuous) definite NP predicate. Consequently, from my analysis, it follows that the *it*-cleft construction can be firmly situated within the language network. In other words, we have an explanation as to why the *it*-cleft construction should exist in the language. Only after examining *it*-clefts in relation to the rest of the grammar are the exceptional or truly construction-specific characteristics isolated. As Goldberg (2003: 118) comments, "a given construction often shares a great deal with other constructions that exist in a language; only certain aspects of its form and function are unaccounted for by other constructions".

Because they don't use inheritance, or look for significant pattern resemblances between constructions, Lambrecht, Davidse and Pavey stipulate enormous amounts of idiosyncratic information. For example, Lambrecht (2001) and Pavey (2004) argue that the *it*-cleft is a noncompositional structure exhibiting complexity mismatch. While Lambrecht (2001) argues that the *it*-cleft expresses a simple semantic proposition via biclausal syntax, Pavey (2004) claims that the semantic "head noun" of the variable expression has no syntactic representation. As I explained in §2.5, noncompositionality is not, contrary to popular belief, a fundamental requirement for positing an independent construction. Most constructional linguists now recognize that the majority of even specialized linguistic patterns are compositional; that is, they can be broken down into units of meaning which can be attributed to individual formal elements. Instead, idiosyncratic information typically pertains to the mapping between form and meaning. In such cases, a unit of form is linked to a unit of meaning with which it is not associated outside of the construction. My analysis of *it*-clefts reflects this aspect of the constructional philosophy. I have presented a largely compositional analysis (with all

but the copula corresponding to a particular unit of semantic meaning) which recognizes the construction-specificity of form-function mappings such as the use of *it* as encoding definite-like quantification rather than denoting a referent.

The so-called "constructional" accounts of Lambrecht (2001), Davidse (2000) and Pavey (2004) therefore capitalize on some of the largely unfounded, popular misconceptions about construction grammar (see Goldberg 1995: 222). They assume that a constructional framework permits and perhaps encourages highly idiosyncratic analyses of noncompositional structures. However, as I have shown, these analyses overlook the main objective of construction grammar: to provide a maximally explanatory account. As a result, I conclude that while the analysis put forward in this thesis conforms to the basic principles of construction grammar, the alternative accounts of Lambrecht, Davidse and Pavey are not actually supported by the constructional framework.

6.3 What are the advantages to a constructional account?

Throughout this chapter, I have noted that construction grammar was developed with specialized linguistic patterns in mind. The *it*-cleft construction is a specialized linguistic pattern. Regardless of how the *it*-cleft's structure is analysed or derived, we are still presented with a nonstandard configuration in which aspects of meaning and use are unpredictable from the independent conventions associated with its component parts. In this section, I compare my constructional analysis to accounts formulated within the generative and transformational traditions. I conclude that the *it*-cleft's properties can be most satisfactorily accounted for and explained within a constructional framework.

One important advantage to examining cleft sentences from a constructional perspective is that it encourages a comprehensive analysis of all aspects of the *it*-cleft's structure and function. As I explained in §2.1, construction grammar does not organize grammatical knowledge into separate components (such as syntax, semantics and so on) which are mediated by highly general linking rules. Instead, mappings between form and meaning are internal to the construction. As a result of this, distinct "levels" of linguistic knowledge cannot be studied in isolation from one another. This means that the object of

linguistic study for the construction grammarian is not only the syntactic configuration, but the construction-specific integration between form and meaning. A constructional approach therefore requires that all of the *it*-cleft's properties are worthy of description and explanation, culminating in a fully comprehensive analysis.

Derivational theories of grammar, on the other hand, tend to focus on form rather than meaning. In this model, a single configuration may have different (underlying) levels of syntactic structure. Therefore, in contrast to the axiom held by most functional linguistics (see especially Bolinger 1977), derivational theories predict that a difference in form is not necessarily accompanied by a difference in meaning. Accounts based on this set of assumptions leave a number of questions unresolved. For example, in §3.2.2, I showed that the inverse analysis of specificational sentences (developed in generative theories such as minimalism and P&P) fails to explain why canonical and inverse word orders have distinct predicational (descriptive) and specificational (identifying) meanings. Likewise, in §3.1, I noted that few proponents of the expletive account of *it*-clefts (which assumes that they are derived from, or otherwise related to, noncopular subject-predicate sentences) tackle the issue of how this construction acquires its pragmatic properties.[6] Finally, the transformational analyses outlined in §4.2 cannot account for the *it*-cleft's construction-specific semantic and discourse-functional properties, which differ from those of their pseudocleft "sources". Consequently, these syntax-based analyses present an incomplete picture of the *it*-cleft and other specificational constructions.

Another advantage to the constructional framework is that the organization of grammatical knowledge as a hierarchical network provides an alternative way of capturing linguistic generalizations. As we have seen, much of the *it*-cleft's structure and use cannot be predicted from highly general rules. However, as shown in §6.2, the constructional framework does not support a wholly idiosyncratic analysis. Likewise, it does not advocate positing a long list of complicated and unmotivated transformations from underlyingly rule-governed structures. Neither of these approaches provides an

[6] Among the expletive accounts, an exception to this generalization is É. Kiss (1998). However, in this case, the pragmatic property of exhaustivity is dealt with as a syntactic phenomenon (see §4.1.2.3).

explanation for why the specialized linguistic pattern, in this case the *it*-cleft, has the particular structure and function that it does. Instead, by representing linguistic knowledge within a hierarchical network, the constructional approach finds explanations at a more local level. For example, when examined in relation to a small family of related constructions, the *it*-cleft quickly loses its mystique and is shown to fit in well with the rest of the language system.

Finally, a constructional approach is useful in the study of the *it*-cleft because it allows for the grammar to contain idiosyncratic information. Obviously, if we label a large proportion of the *it*-cleft's properties as construction-specific, then our analysis loses its explanatory value. However, construction grammar simply states that while most properties of a given construction are inherited from more general patterns of correspondence, it is normal for some aspects of its form or function to be construction-specific. As I explained in §6.1, by factoring out some idiosyncratic properties, our analysis of *it*-clefts becomes far more straightforward, allowing the relevant generalizations to be made. Furthermore, by recognizing the existence of idiosyncratic properties, the constructional framework asks us to explain or provide motivation for them. This represents an important advantage over theories that do not explicitly predict or allow idiosyncratic information to enter into the grammar. For such accounts, the only options are either to ignore construction-specific properties completely, thereby presenting an incomplete analysis of the construction, or to provide a complex and elaborate derivational story in which the construction is ultimately traced back to an underlyingly rule-governed and predictable structure.

The way that construction grammar provides motivation for idiosyncratic properties is by looking for an explanation outside of the synchronic language network. As a usage-based model, the version of construction grammar adopted in this thesis is designed to intersect with and work alongside theories of language learning and language change[7]. In the remainder of this thesis, I examine whether historical evidence can provide an explanation for some of the construction-specific properties that *it*-clefts

[7] As I explained in chapter 2, this thesis makes use of a usage-based constructional framework which argues for a default inheritance model of language in which information is stored redundantly. However, not all versions of construction grammar share these properties.

display. *Contra* Roberts and Roussou (2003), I go on to show that construction grammar lends itself to integration with grammaticalization theory better than the minimalist framework.

In conclusion then, the constructional approach encourages an all-encompassing, explanatory analysis which highlights the maximum level of correspondence with other constructions. At the same time, construction grammar prevents the stipulation of over-complex and *ad hoc* derivations by allowing for idiosyncratic information to enter into the grammar. Motivation for specialized linguistic patterns is provided either by the structured inventory of the language system or by language external factors. Finally, and of particular importance to the analysis of *it*-clefts argued for in this thesis, construction grammar has a greater level of diachronic applicability than alternative compositional (i.e. non-constructional) frameworks.

6.4 A hierarchy of constructions

The topic of this chapter has been the framework in which my analysis is conducted. I have explained what is constructional about my approach, how it compares to other so-called "constructional" accounts of *it*-clefts, and why construction grammar is particularly suited to the study of such specialized linguistic patterns. In this section, we come to the end of my synchronic account of *it*-clefts. This is as far as most authors writing about cleft sentences go. However, as I have noted in §6.3, one of the advantages to the constructional framework is that it encourages us to look for explanations further afield and provides us with the tools to do so. As a result, in chapters 7 and 8, I turn my attention to the historical development of the *it*-cleft construction. As a precursor to this, in this section, I highlight some of the questions that my synchronic account has yet to provide answers for. These will form the basis of my diachronic investigation. However, first, I review the inheritance links which have formed the basis of my analysis so far and present this family of constructions within a hierarchical network (see Figure 6.3).

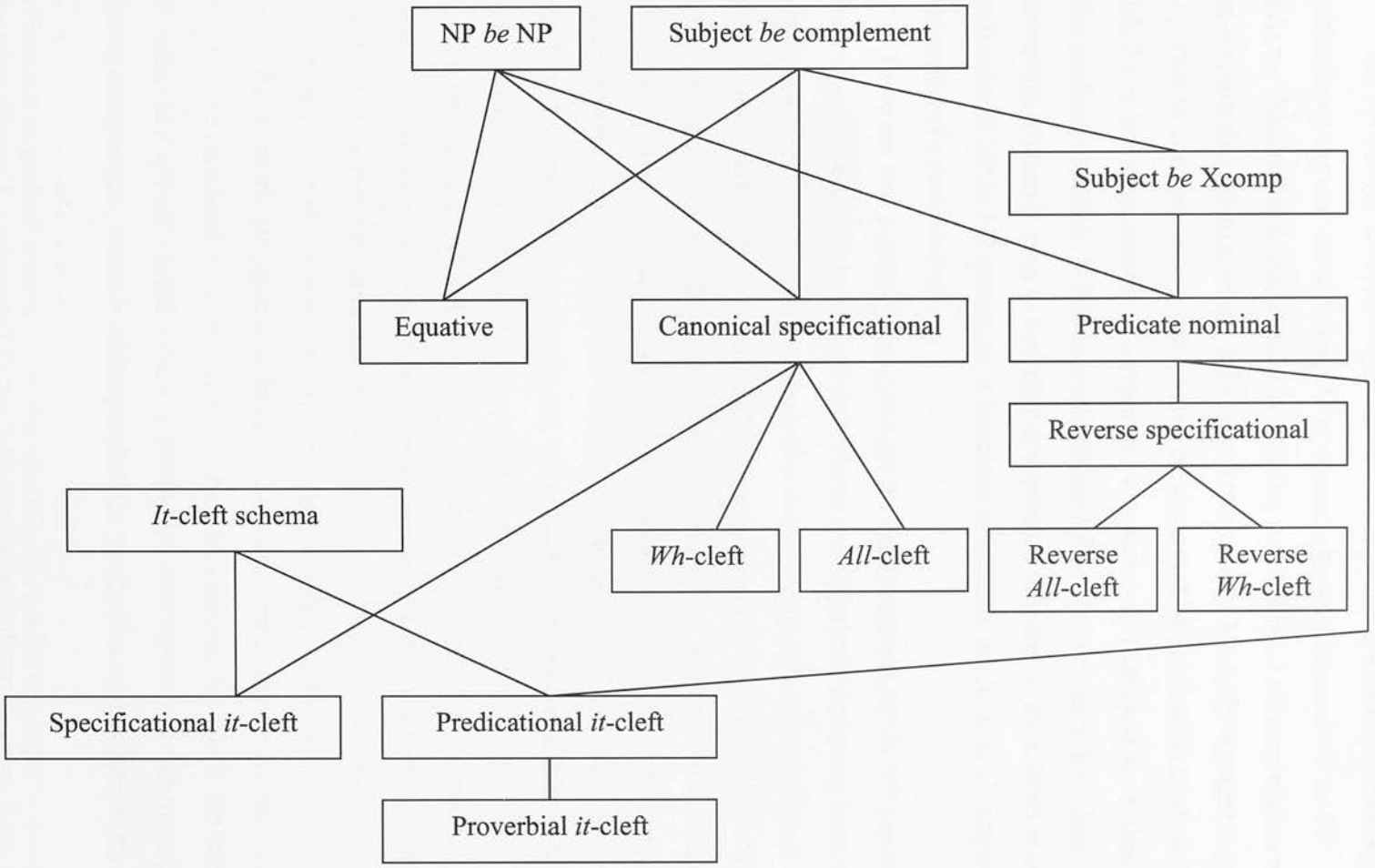


Figure 6.3 *An interim inheritance hierarchy of* NP be NP *constructions*

My synchronic analysis began with the premise that *it*-clefts are a variety of specificational copular construction. After examining specificational *NP be NP* sentences, I claimed that specificational meaning results from a reinterpretation of the class-membership relation which is found in predicational *NP be NP* sentences (see §3.3). That is, both sentence types involve the same type of predication relation, in which the referring expression is categorized as a member of the class or set described by the predicate nominal. While in predicational *NP be NP* sentences, this class-membership relation is used to **describe** or attribute a property to the subject referent, in specificational *NP be NP* sentences, it functions to **identify** the (typically complete) membership of a restricted set.

There are two possible configurations for specificational *NP be NP* sentences: canonical and reverse. I have argued that reverse specificational sentences form a sub-construction of the larger predicate nominal construction (see Figure 6.3).[8] Such sentences inherit all of the properties of the more general predicational *NP be NP* schema, including syntactic elements, semantic components and the symbolic links between them. However, the reverse specificational construction is a more specialized linguistic pattern, stipulating that the semantic predicate is a definite noun phrase (providing the prerequisite restrictive set) and exhibiting an additional information-structural requirement (such that the precopular referent is in focus).

In canonical specificational sentences, the restricted set (the predicate) is presented before the identification of any of its members. I have claimed that this construction involves mismatch, overriding the pervasive generalization that the syntactic subject functions as a semantic argument. In other words, the symbolic links between the syntactic elements and the semantic components are not inherited, but are specific to the canonical specificational construction. I account for this in the following way. Canonical specificational sentences exhibit the same syntactic structure as the equative construction, which is underspecified for predicative complement status.

---

[8] In Figure 6.3, the predicate nominal construction inherits its structure from the *subject be Xcomp* construction. 'Xcomp' is a label, borrowed from LFG, meaning predicative complement. In this hierarchy, I have suggested that the *subject be Xcomp* structure is a less schematic variant of the underspecified *subject be complement* construction.

Instead, the nominal predication relation comes about compositionally as a result of the inherent semantics of definite noun phrases. According to a Russellian-Gricean analysis, definite noun phrases are semantically quantificational, involving the universal (proportional) quantification of a restrictive and existentially presupposed set. Consequently, when a definite noun phrase is compared with a referring expression, a class-membership predication relation ensues[9], resulting in a specificational sentence.

In §4.1, I claimed that *it*-clefts with a specificational (or identifying) meaning are a subtype of specificational copular construction.[10] In order to maximize motivation for the *it*-cleft, I advocated a discontinuous constituent analysis. From this perspective, the *it*-cleft is reduced to an *NP be NP* sentence, displaying a class-membership predication relation. This, in turn, explains why *it*-clefts have a specificational function. In addition, I have shown that a number of the *it*-cleft's pragmatic properties, such as a fixed focus position, existentiality, exhaustiveness, contrast and givenness are inherited from the wider specificational schema and can often be explained as originating from the inherent semantics of definite noun phrases. Likewise, this analysis enables the cleft clause to be unequivocally classified as restrictive relative (which forms part of the discontinuous definite description), thereby reducing the number of construction-specific properties that the *it*-cleft displays.

As well as having a specificational function, *it*-clefts can also occur with a predicational (descriptive) meaning. Given my analysis of *it*-clefts as akin to simple *NP be NP* sentences, this is to be expected. Predicational *it*-clefts have the same syntax, semantics and information structure as specificational *it*-clefts. This is illustrated in Figure 6.3, with predicational and specificational *it*-clefts inheriting construction-

---

[9] I have explained, in §3.3.4, that since indefinite noun phrases involve cardinal, rather than proportional quantification, they denote an individual member, rather than a set or subset. As a result, they very rarely occur as the predicate of a specificational sentence and are permitted (via coercion) only if the indefinite NP assumes the qualities of a definite description. However, since this coerced interpretation creates its own set of discourse functions, specificational sentences with indefinite NP predicates could perhaps form a distinct sub-construction in the speaker's mind. However, this is not represented in the over-simplified hierarchy given in Figure 6.3. Likewise, the *there*-cleft construction (which is a non-exhaustive variant of the *it*-cleft) is also absent.

[10] As shown in Figure 6.3, *wh*-clefts and *all*-clefts are also subtypes of the specificational copular construction. *Th*-clefts are not recognized as a distinct construction in this hierarchy since they cannot be clearly separated from "noncleft" *NP be NP* sentences (see §4.1.1).

specific properties from a common *it*-cleft abstraction. However, as I explained in §5.3, these sentence types differ in their internal form-function mappings; in predicational *it*-clefts the syntactic subject functions as a referential argument expression and the postcopular NP is the semantic predicate. As shown in Figure 6.3, this property is inherited from the predicate nominal construction.

These constructional correspondences and inheritance relations are sketched in the constructional taxonomy illustrated in Figure 6.3. However, this temporary network is incomplete and leaves a number of questions unresolved. For example, it does not tell us how the *it*-cleft's semi-substantive and construction-specific structure is motivated. Why does the constituent *it* not function as a fully referential pronoun? How can a pronoun be modified by a restrictive relative clause to form a definite-like description? How is the extraposition of the restrictive relative clause motivated? Why do we not have non-extraposed forms of the *it*-cleft in which *it* and the relative clause form a "continuous constituent"? In addition, the constructional taxonomy only deals with structures containing NP foci. How do non-NP specificational *it*-clefts relate to this family of constructions and how are they motivated? Finally, we have not yet placed informative-presupposition *it*-clefts, which occur with new information in the relative clause, within this hierarchy. As I explained in §5.5, this property is not shared with other specificational copular constructions. So how is it motivated?

These questions form the basis of the diachronic investigation undertaken over the next two chapters. Here, I show that historical evidence can help to provide motivation for the *it*-cleft in two different ways. First, it provides us with a picture of the language system at earlier periods of the language. As we have seen, some of the *it*-cleft's formal properties cannot be traced back to more general constructions in the present-day language system. However, this does not mean that these properties were not, at one time, inherited. As I go on to show, historical investigation can identify the existence of formally related constructions at earlier periods of the English language which can help to explain the *it*-cleft's unusual configuration. In addition, diachronic corpus research shows how the *it*-cleft has developed over time. While the *it*-cleft may be formally and functionally related to a family of constructions, language change does

not always happen on such a grand scale. As a result, historical data can show how construction-specific properties develop over time. Combined with a theory of language change, such as grammaticalization, we can begin to get a picture of why the *it*-cleft has developed these idiosyncratic properties and what effect this has on its modern day distributional potential.

7. A HISTORICAL INVESTIGATION OF *IT*-CLEFTS

In chapter 6, I showed that the analysis of *it*-clefts argued for in this thesis maximizes motivation for the construction; in other words, it provides an explanation as to why this particular construction with these particular properties should exist in the language. In this account, specificational and predicational *it*-clefts inherit many of their properties from more basic constructions including the (canonical) specificational copular construction, the predicate nominal construction and the definite noun phrase construction. In this way, a large number of the *it*-cleft's formal and functional properties are provided with an explanation. Nevertheless, on this account, there are still aspects of the *it*-cleft's structure and use that remain unexplained.

In this chapter, I show how, once we look beyond inheritance from the present-day language system, we can find an explanation for many of the *it*-cleft's construction-specific properties. Here, I undertake a historical investigation into the *it*-cleft's origin and subsequent diachronic development, showing how and why the *it*-cleft acquired its idiosyncratic characteristics. The historical evidence integrates in several important ways with the synchronic *it*-cleft analysis developed over previous chapters. First, my extraposition-from-NP account of *it*-clefts captures the diachronic facts much more straightforwardly than an expletive analysis. As a result, I am able to provide a much simpler story of the historical development of the *it*-cleft than Ball (1991, 1994a), whose work dominates the somewhat limited literature on this topic. This, in turn, provides additional support for an extraposition-from-NP analysis of *it*-clefts. Together, my synchronic and diachronic analyses add up to a maximally explanatory account of the *it*-cleft construction.

There are two different ways in which I make use of historical evidence in this chapter. First, in §7.1, I examine *it*-cleft examples from earlier periods of English in relation to the synchronic language system of the time. This allows me to identify some inheritance relations that are no longer productive in the constructional hierarchy. I show that while aspects of the *it*-cleft's structure are left unexplained when examined in relation to the present-day language system, they were at one time motivated by

constructions that have subsequently fallen out of use or which have undergone changes to their formal properties. In this section, I make use of and reanalyse Ball's (1991) Old English and Early Middle English data. I find that *it*-clefts were, from the very beginning, more amenable to the extraposition-from-NP analysis argued for in this thesis than the expletive analysis adopted by Ball (1991). This leads me to propose an alternative origin story for the specificational *it*-cleft, dating back to Old English. I show that the *it*-cleft construction was, at this time, fully motivated by inheritance from the language system.

The second way that I make use of historical evidence is by examining the effects of language change on the *it*-cleft construction. As I explained in §6.1, in construction grammar, motivation for a construction can be provided not only by inheritance relations that are internal to the structured inventory of the language, but also by factors external to the grammar, including processing, acquisition and language change (see Goldberg 2003: 121). In §7.2, I undertake a diachronic investigation into the development of the *it*-cleft construction throughout Late Middle English and Early Modern English. Here, I show that over time the *it*-cleft construction gradually sanctions a wider variety of instances which override more general patterns of correspondence. As I go on to explain, these changes are in keeping with the unidirectional tendencies of 'grammatical constructionalization' which have been outlined within grammaticalization theory. I conclude that the *it*-cleft's emergent idiosyncratic properties are motivated by, and obtain an explanation from, general principles of language change.

7.1 The *it*-cleft and the language system of earlier periods of English

Throughout this thesis, I have argued for an analysis which treats the *it*-cleft as a subtype of *NP be NP* sentence. I have shown that once we analyse the initial *it* and the cleft clause as a discontinuous (restrictively modified) definite description, we are able to explain the specifying (or, for predicational clefts, the ascriptive) function of the construction as well as its shared behaviour with definite noun phrases and restrictive relative clauses. However, there are aspects of the *it*-cleft's structure that remain a puzzle on this account. How is it possible that a restrictive relative clause modifies a

pronominal element, such as the initial *it*? How is this motivated? If the initial *it* and the cleft clause form a semantic constituent, then why can they never be adjacent to one another? Why must the cleft clause always occur in sentence-final position and what motivates this extraposition-from-NP? Finally, why doesn't the embedded verb in the relative clause always show number agreement with the initial *it*?

In this section, I show that these construction-specific formal properties have their sources in constructions that we haven't yet considered; I argue that they were originally inherited from syntactic configurations which no longer exist (or are no longer productive) in the language system. I show that while the *it*-cleft's structure has remained largely unchanged since Old English, what has changed is the language system in which the *it*-cleft construction is situated. This idea, that irregularity is sometimes the relic of fossilized historical regularity, is well-understood. For example, irregular past tense forms such as *stand/stood* are the entrenched relics of a once productive pattern of strong verbs. Likewise, as I explained in §2.5, idioms such as *kith and kin* (meaning *friends and family*) and *with might and main* (meaning *with a lot of strength*) contain words which are no longer found outside of these formulaic expressions (see Fillmore, Kay and O'Connor 1988).

In §7.1.1, I provide evidence that the discontinuous constituent of the *it*-cleft is an instance of a once productive construction in which pronouns are modified by restrictive relative clauses (the 'determinative pronoun construction'). As part of this discussion, I examine early *it*-cleft examples taken from Ball (1991). While these examples pose a problem for Ball's expletive analysis, they are accommodated neatly by my extraposition-from-NP account and provide unequivocal support for it. This suggests a much simpler origin story for the specificational *it*-cleft than that argued for by Ball (1991). In §7.1.2, I show that the sentence-final position of the relative clause was the norm in Old English, arguing that the *it*-cleft configuration represents a reflex of the older pattern. Finally, in §7.1.3, I claim that the cleft pronoun *it* has more in common semantically with the Old and Early Middle English lexeme *it* than with the present-day pronoun. I claim that with this analysis in place, we can better understand the *it*-cleft's unusual agreement patterns. A summary of my findings is given in §7.1.4.

7.1.1 How can a pronoun be restrictively modified?

In my account of the *it*-cleft construction, the cleft clause is analysed as a restrictive relative which modifies the initial *it*. However, a potential problem with this account is that relative clauses do not normally modify pronominal elements. Nevertheless, at earlier periods of the language, pronouns could function as the antecedents to restrictive relative clauses. Although no longer productive, such patterns exist in formulaic sayings such as (1).[1] Here, the pronoun *he* and the following relative clause function as a definite description, akin to *the one who laughs last.*

(1) He who laughs last laughs longest

Following Declerck (1988) and Ball (1991) among others, I label the pronouns which are modified in this way 'determinative pronouns'. In what follows, I use historical evidence to build up a case for analysing the cleft *it* as a determinative pronoun. I argue that although the restrictive modification of *it* is now an entrenched and idiosyncratic part of the *it*-cleft schema, it was, at one time, a property inherited from the then productive 'determinative pronoun construction'.[2]

Discussing proverbial predicational *it*-clefts, Declerck (1988) provides evidence for the claim that, in this particular type of *it*-cleft, the initial *it* is a determinative pronoun, modified by a restrictive relative clause. For one thing, Declerck (1988: 155) notes that the initial *it* in these examples can sometimes be replaced by other determinative pronouns, like *he* or *she*. Such examples, including (3) below, show that it is possible for the restrictive relative clause which modifies the determinative pronoun to be located in an extraposed position. This provides support for our extraposition-from-NP analysis of the corresponding *it*-cleft given in (2).

---

[1] In addition, indefinite pronouns can be restrictively modified in contemporary English. For example, in the NP *someone who lives here*, the relative clause restricts the very general type specification *someone*. Furthermore, for some speakers, definite pronouns can be followed by reduced relatives in expressions such as *them round the corner* and *him next door*.

[2] Curme (1931) analyses determinative pronouns differently from other authors. He suggests that the clausal component in *it*-clefts and other determinative pronoun constructions is a "subject clause" rather than a restrictive relative. On this account, the cleft *it* and other determinatives serve as "anticipatory subject[s], pointing to the following subject clause" (Curme 1931: 188).

(2) It is a fortunate voyager who finds many friends

(3) ...and the best that we find in these travels is an honest friend. He is a fortunate voyager who finds many. (Kruisinga 1932: 505; cited in Declerck 1988: 155)

Furthermore, Declerck observes that, along with *he* and *she*, the pronoun *it* could at one time function as a determinative pronoun in contexts outside of the cleft construction. As noted by Poutsma (1928: 730) (cited in Declerck (1988: 156)), the pronoun *it* in the following example from Shakespeare's *Twelfth Night* is "used as a kind of determinative".

(4) I would have men of such constancy put to sea, that there business might be everything and there intent everywhere, for that's it that always makes a good voyage of nothing. (*Twelfth Night;* Act 2 Scene 4)

In this predicate nominal construction, the restrictive relative clause immediately follows the pronoun *it*. Together these elements make up a definite-like description (with the meaning *the thing/quality that makes a good voyage come to nothing*) which is predicated of the anaphoric subject *that*.

Of course, in this example, the postcopular noun phrase is predicated of an inanimate referent. If we want to provide *it*-clefts such as (2) with a determinative pronoun analysis, then we need to provide an explanation for the fact that the non-human *it* is the head of a noun phrase which has a human referent, in this case *the voyager that finds many friends*. In response to this challenge, Declerck (1988: 156) notes that in former times, *it* and *he/she* were largely interchangeable. He cites the following example from Visser (1970: 41) in which the pronoun *it* is used as the subject of a predicational sentence to refer to what must be a human referent (*a bishop*).

(5) Hit is an biscop (c971, *The Bickling Homilies*, 43.33)

Taken together, this historical evidence shows that (a) in proverbial expressions, pronouns other than *it* can be modified by restrictive relative clauses, which are sometimes extraposed; (b) *it* could at one time function as a determinative pronoun in noncleft contexts; and (c) although pronominal *it* can now only refer to non-human entities, at earlier stages of English, *it* could have a human referent. On the basis of this evidence, Declerck concludes that the initial *it* in proverbial predicational clefts is a determinative pronoun. In such sentences then, the pronoun *it* and the cleft clause form a discontinuous definite description which functions as the referring expression within a predicate nominal construction (see §5.2). As Declerck (1988: 157) notes, a determinative pronoun analysis for these examples is supported by the fact that, as proverbs, they "are reminiscent of an earlier stage of the language".

Declerck (1988) therefore limits his determinative pronoun analysis to proverbial predicational clefts. However, in the remainder of this section, I show that historical evidence also provides support for a determinative pronoun (or extraposition-from-NP) analysis of both specificational and non-proverbial predicational *it*-clefts. I conclude that it is not only the individual proverbial tokens which are entrenched in the language system; in fact, the entire *it*-cleft schema has remained largely unchanged since Old English, and retains residual patterns which are no longer productive outside of the cleft construction. In what follows, I discuss early examples of *it*-clefts taken from Ball's (1991) data set but, crucially, I do not always agree with her analyses.

As part of her investigation into the historical development of the *it*-cleft, Ball (1991) identifies an Old English cleft configuration of the form ÞIS/ÞÆT/HIT BEON NP REL-CLAUSE. The Old English cleft data is extremely instructive, since, at this stage of the language, all nouns are marked for gender. We therefore have an additional means with which to identify the structure of these early cleft sentences. Tellingly, in Ball's ÞIS/ÞÆT/HIT BEON NP REL-CLAUSE examples, the relative pronoun shows gender agreement with the initial pronoun (*hit, þæt* or *þis*) rather than with the postcopular NP. From this, we can conclude that these Old English cleft tokens require a determinative pronoun analysis, in which the relative clause restrictively modifies the subject pronoun.

For example, in the Old English predicational *þæt*-cleft given here as (6), the neuter relative pronoun *þæt* agrees with the subject pronoun rather than with the masculine noun *journey*. Ball (1991: 59) notes that "The gender of the relative pronoun suggests that [the initial] *þæt* is the head of the relative clause". She therefore provides a determinative pronoun (extraposition-from-NP) analysis for this example.

(6) *wiste his fingra geweald on grames grapum. Þæt wæs geocor sið*
DEM-**n**.n.s was grievous journey-**m**.
þæt se hearmscaþa to Heorute ateah! (*Beowulf*, 764b-766)
REL-**n**.s. DEM-m.n.s despoiler to Heorot took (Matsunami 1961: 7)
'[he] knew his fingers' power to be in a hateful grip. That was a painful journey that the loathsome despoiler had made to Heorot.' (Ball 1991: 35, 54)

Adopting Culicover and Rochemont's (1990) account of 20thC English extraposed relative clauses, Ball (1991: 63) claims that, in Old English predicational clefts, the relative clause is base-generated in an extraposed position adjoined to IP and is coindexed with the neuter subject pronoun (*hit, þæt* or *þis*).[3] Ball provides present-day predicational *it*-clefts and demonstrative clefts with the same analysis. She concludes that the structure of the predicational *it*-cleft has remained relatively unchanged since Old English.

The same gender agreement patterns are found in Ball's non-predicational ÞIS/ÞÆT/HIT BEON NP REL-CLAUSE tokens, such as the *hit*-cleft given here as (7). As is shown in Mitchell's (1985: 102) gloss, neuter *þæt* does not agree with either *Petrus* or *engel*, both of which are masculine. This leads both Mitchell (1985: 102) and Ball (1991: 67) to conclude that the initial pronoun *hit* is the antecedent to the relative clause.

[3] However, Ball (1991:55) notes that the meaning of (6) is not '*that which the loathsome despoiler had made to Heorot was a painful journey*'. As with present-day predicational clefts, the postcopular noun provides the hearer/reader with the required interpretation for the underspecified subject pronoun (see §5.2.2). The meaning of (6) is therefore '*the journey that the loathsome despoiler had made to Heorot was a painful journey*'.

(7) *þa cwædon þa geleafullan, 'Nis hit na Petrus þæt þær cnucað, ac is*
Not-is it-**n**. not Peter-**m**. REL-**n**. there knocks but is
his ængel.' (*Ælfric, Catholic Homilies*, vol. I, 517-18.1)
his angel-**m**. (Mitchell 1985: 102)
'Then the faithful said: It isn't Peter who is knocking there, but his angel.'
(Ball 1991: 39)

In this example, the faithful mistake St. Peter, who they believe to be in prison, for his angel; that is, they (wrongly) identify *his angel* as *the one who is knocking there*. I suggest that these non-predicational Old English *hit*-clefts represent early instances of the specificational *it*-cleft construction. Such examples provide unequivocal evidence that, from the very beginning, the specificational *it*-cleft requires an extraposition-from-NP analysis, therefore supporting one of the main claims of this thesis.

The Old English cleft data therefore shows that the structure of present-day specificational and predicational *it*-clefts (and demonstrative clefts) derives from a period of the language when determinative pronouns were much more common.[4] Originally then, the *it*-cleft inherited its structure from the then productive determinative pronoun construction, in which pronouns function as the antecedents to restrictive relative clauses. While the basic structure of the *it*-cleft has remained unchanged, determinative pronouns are now largely restricted to proverbial expressions, such as (1) above. From this, we can conclude that although the modification of the initial *it* by the cleft clause is now a construction-specific property, it represents the entrenchment of a schema which was at one time motivated by the language system.

However, this simple diachronic story is dependent upon an extraposition-from-NP analysis of the present-day *it*-cleft. Ball (1991), who adopts an expletive analysis of

[4] Tellingly, restrictively modified pronouns could at one time serve as the definite NP predicate of other types of specificational sentence. For instance, in the following Middle English token, the pronoun *he* is modified by the adjacent relative clause, forming a definite description which is predicated of the referential subject. In this example, the referent *ye* is identified as *the one* (or '*he*') *that killed Marhaus*. This type of reverse specificational sentence was later replaced by *th*-clefts introduced by *the one*.

(i) *Truly, said Bleoberys, I am ryght gladde of you, for ye are he that slewe Marhaus the knight....*
(1485 Malory's *Le Morte d'Arthur* (Cx), 217.2; cited in Ball 1991: 231)

the specificational *it*-cleft, presents a very different account of the diachronic data. In what follows, I outline Ball's analysis and the reasoning behind it. I show that an expletive *it*-cleft analysis is unable to accommodate all of the diachronic facts. This leads Ball to put forward an unnecessarily complex and fundamentally flawed account of the early *it*-cleft data. Again, this discussion provides support for the analysis of specificational *it*-clefts argued for in this thesis and shows how my diachronic story improves on the existing literature.

From the outset, Ball assumes an expletive analysis of the specificational *it*-cleft, in which the cleft pronoun is a dummy element and the cleft clause is a type of headless restrictive relative. However, the Old English cleft data presents a problem for Ball since her expletive analysis cannot accommodate it. As shown above, the gender agreement pattern found in these tokens calls for an extraposition-from-NP account. Ball (1991: 39-40) comments that, given her expletive analysis of the present-day specificational *it*-cleft, "This is a surprising fact: if these are true clefts, in which the focus is the logical antecedent of the relative pronoun, we should expect gender agreement with the focus." In order to get around this problem, Ball claims that the non-predicational *hit*-cleft examples are not in fact early instances of the specificational *it*-cleft; instead, they make up an entirely different construction.

Ball attempts to substantiate this claim by invoking Higgins' (1979) distinction between specificational and identificational copular sentences. For Higgins, examples such as (8) have two possible readings. On the specificational reading, the person *Mary Gray* is identified as matching the description *the girl who helps us on Fridays*. On the identificational reading, we are told the name (*Mary Gray*) of the person referred to by the definite NP *the girl who helps us on Fridays*.

(8) The girl who helps us on Fridays is Mary GRAY (Higgins 1979: 265)

We have seen that, in specificational *NP be NP* sentences, the syntactic subject is a definite NP predicate and the postcopular NP is a referring expression (see §3.3).[5] However, for identificational copular sentences, Higgins (1979: 264) notes that the subject NP is a referring expression and the postcopular NP is 'identificational'; in other words, it is a label or naming device. As I explained in §3.2.3, the category of identificational copular sentences is not well-defined and Mikkelsen (2005) has argued that Higgins' examples can be accommodated into the tripartite taxonomy of predicational, equative and specificational sentences.[6]

Despite this, Ball sets up a categorial distinction between specificational and identificational meaning, arguing that her Old English non-predicational *hit*-clefts are actually identificational, rather than specificational, copular sentences. For example, she claims that in (7) above, "there is an otherwise established entity in the context [*the one knocking at the door*] which lacks only a name" (Ball 1991: 64). However, this cannot be right. In (7), what is at issue is not the name of the referent (as *Peter* or *his angel*) but the referent itself. This example clearly conforms to our understanding of specificational meaning as involving the identification of the referent which uniquely matches the definite description. Indeed, Ball (1991: 66) herself acknowledges that the Old English *hit*-clefts allow a specificational reading; she admits that "these tokens could also be read as specificational" and concludes that "There is a fine line between identification and specification".

Tellingly, Ball only makes use of the specificational/identificational distinction when discussing the Old English data. For subsequent periods in the history of the *it*-cleft, Ball finds it difficult to implement this distinction as a means of classification and consequently abandons it. She notes that "Specificational and identificational *it*-clefts are classed together here because of the difficulty of reliably distinguishing them…While the distinct semantics of the predicational cleft continue to set it apart, the

---

[5] However, for Higgins (1979), specificational sentences contain 'superscriptional' subjects and 'specificational' complements. See §3.2.3 for a review of Higgins' proposal.

[6] Assuming a tripartite taxonomy of copular sentences, the identificational reading of (8), involving a referential subject and a postcopular naming device, must, in fact, be predicational. Such sentences contain a referential subject and a predicative complement. On this account, the identificational reading of (8) is an elliptical version of *the girl who helps us on Fridays is* ***called*** *Mary Gray*.

other two types are close both in interpretation and function" (Ball 1991: 220). This seriously undermines Ball's claim that together predicational and identificational clefts form a separate construction, with a distinct structure, from the specificational *it*-cleft. It seems to me that if, as Ball suggests, there is a real distinction between specificational *it*-clefts, which require an expletive analysis, and identificational *it*-clefts, which require an extraposition-from-NP analysis, then Ball would make more of an effort to separate the progress of these two very different constructions.

Ball's inconsistent use of the ill-defined distinction between specificational and identificational meaning is therefore, quite obviously, just a (very unsuccessful) means of getting around the problem that her expletive analysis cannot accommodate the Old English data. It is only when the system of gender marking (and therefore gender agreement) breaks down that Ball can argue for an expletive analysis of the specificational *it*-cleft examples. This leads Ball to argue that the origin of the specificational *it*-cleft is an Early Middle English development. She provides the following example from *The South English Legendary*, which she argues is the first specificational *hit*-cleft in the corpus.

(9) 'A-bidez,' quath þis holie man: 'ore louerd is guod and freo.
þe deuel it is þat bringuth þis wedur…'
(1280-90, *The South English Legendary*, 63 (St. Edmund), 370)
[As St. Edmund was preaching, it became overcast, a terrible wind began to blow, and it grew dark. People began to leave.] 'Stay, said this holy man, our Lord is good and free. The devil it is that brings this weather…' (Ball 1991: 158)

Rather than developing from the Old English "identificational" examples, Ball claims that the primary source construction for this new sentence type was the Old English NP BEON REL-CLAUSE configuration, an example of which is given in (10) below. Such sentences can be translated as present-day *it*-clefts but, as Ball (1991: 51) notes, they are formally akin to reverse pseudoclefts, since the focal NP is the grammatical

subject.[7] The only difference is that, in these tokens, the predicate complement is a headless relative ((*the one*) *that glorifies me*). In Old English, relative clauses without overt heads often occurred as the complement of *beon* (Ball 1991: 27).

(10) *...min fæder is þe me wuldrað* (*Ælfric, Catholic Homilies*, vol. II, 234.4)
...my father is that me glorifies
'It is my father that glorifies me' (Ball 1991: 27)

Ball suggests that, during the Early Middle English period, some of these examples acquired expletive subjects, resulting in the origin of the specificational *it*-cleft. She claims that "the focused subject...becomes the predicate complement, and dummy *hit* appears in its place. The resulting structure is superficially similar to [the OE identificational *hit*-cleft] except that the complement is within the VP, and this *hit* is expletive" (Ball 1990: 68).

Ball's adoption of the expletive account of specificational *it*-clefts therefore has a significant effect on the way that she categorizes and analyses the historical data. It leads her to discount important Old English tokens and to provide an unnecessarily complex story for the origin of the specificational *it*-cleft. In contrast, the extraposition-from-NP analysis argued for in this thesis can accommodate all of the historical data and suggests a much more straightforward diachronic story. I have shown that the present-day *it*-cleft (both specificational and predicational subtypes) can be traced back to Old English. During this period, the *it*-cleft's extraposition-from-NP structure was motivated by the then productive determinative pronoun construction.

7.1.2 Why is the restrictive relative clause extraposed?

In the previous section, I showed that the *it*-cleft's structure has remained relatively unchanged since Old English. Of particular note, is the fact that, from the very beginning, the restrictive relative clause in *it*-clefts is extraposed and is never found in a

[7] In support of this claim, Ball (1991: 27) notes that these tokens are often translations of Latin pseudoclefts. For example, (10) is an English translation of the Latin original *...est Pater meus, qui glorificat me.*

position adjacent to the antecedent *it*. As I explained in §4.3, the fact that the cleft clause always follows the clefted constituent, rather than the initial *it* has been regarded by many as a valid argument against an extraposition-from-NP analysis (see for example Jespersen 1937: 84-85). However, in this section, I show that there is a perfectly reasonable, historical explanation for the *it*-cleft's extraposed structure.

Since the particular *it*-cleft analysis argued for in this thesis is non-derivational, it does not require a non-extraposed source construction. Nevertheless, the question remains, why is the cleft clause never adjoined to the antecedent *it*? The Old English *hit*-clefts seem to provide us with a suitable answer. Although the extraposed position of the restrictive relative is atypical in relation to the rest of the present-day language system, relative clauses were often found sentence-finally in Old English. As Ball (1991: 60) notes, "it is commonly held that OE relative clauses originated as paratactic structures, which gradually became more integrated with the sentences to which they were attached" (see for example O'Neil 1977). It seems then that this now idiosyncratic property of the *it*-cleft construction was at one time motivated by the general behaviour of Old English relative clauses. Ball (1991: 60) comes to this same conclusion for the predicational *it*-cleft; she suggests that unlike other relative clauses, the cleft clause "never came together with its antecedent, but remained in fixed clause-final position."

The obvious question then, is why didn't changes to the wider language system have similar consequences for the *it*-cleft construction? One reason seems to be that while the extraposition of the cleft clause is no longer motivated by highly general linguistic patterns, it is motivated by information structure principles, such as the tendency for heavy constituents to occur near the end of the clause. Findings from Prince's (1978: 886) corpus study suggest that, on average, "the presupposed string [the cleft clause] is nearly twice as long as the focused string". This, accompanied by the fact that new or focal information is placed within the cognitively preferred clause-final position, makes the cleft structure well suited to linguistic generalizations brought about by processing demands.

In addition, prosodic factors may well provide another reason for the obligatory sentence-final position of the cleft clause. As shown in example (4) above, the

determinative *it* in noncleft constructions could be modified by an adjoining relative clause. However, in this case, the definite-like description is in the postcopular predicative position rather than in initial subject position. Tellingly, Ball (1991: 263) notes that "for whatever reason, *it* + relative clause is not attested in first position". This suggests that the reason why the cleft clause cannot occur immediately after initial *it* is part of a more general phenomenon affecting the determinative pronoun. Bolinger (1977) provides us with a possible explanation for this restriction, attributing it to the prosody of English. He notes that "*It* is normally stressless, but in initial position followed by an obligatory stressless *that* it would have to be stressed" (Bolinger 1977: 76).

We can conclude then that while the extraposed position of the cleft clause is atypical in relation to the present-day language system, this was not always the case. At earlier stages of English, the behaviour of the cleft clause was consistent with that of other restrictive relatives, which frequently occurred in sentence-final position. As the external structure of relative clauses has changed, leading to more integration between the relative and its antecedent, the *it*-cleft has remained the same. Now, the extraposition of the cleft clause is motivated purely by information structure principles and/or prosodic factors.

However, although the extraposed relationship between the cleft clause and its antecedent is now specific to the *it*-cleft, the internal structure of the cleft clause is nevertheless inherited from the restrictive relative construction. Again, this is supported by historical evidence. Ball (1994b) undertakes a comparative study of the relative markers in NP-focus *it*-clefts and restrictive relative clauses. She finds that the cleft clause and other restrictive relatives have undergone parallel diachronic developments.

Focusing on clauses in which the gap functions as the subject of the embedded verb, Ball shows that in both *it*-clefts and restrictive relatives, *that* is the preferred complementizer up until the 18th century, when there is a sharp increase in the use of *wh*-pronouns. However, Ball notes that while the pronoun *which* undergoes a rise in the 18th century, it never achieves the same high frequency as *who*. Ball interprets this data as evidence of a paradigmatic shift in the system of relative markers (see also Ball

1996). She notes that during the 17th century, *who* replaces personal *which*, leading to a distinction in the restrictive relative paradigm between personal and non-personal antecedents. This is followed by the assignment of *that* to the non-personal category. As a result of the decline in *that* with personal antecedents, *who* has become the dominant form for cleft clauses and restrictive relatives with personal subject antecedents, at least in standard written English. On the other hand, non-personal *which* remains in competition with *that* and so is not subject to the same rate of increase.

Ball argues that *it*-clefts and restrictive relative clauses undergo these same changes at roughly the same time and at the same rate. She notes that "the rate of decrease in *that* with personal antecedents is not significantly different in clefts and restrictive relatives, a finding which supports the hypothesis that cleft complements and restrictive relative clauses share a syntactic structure" (Ball 1994b: 196). As a result, the historical evidence suggests that while the extraposition of the cleft clause is reminiscent of an earlier stage of the language, it is nevertheless a fully-fledged restrictive relative which inherits its internal structure from the more productive construction.

7.1.3 Why does the it*-cleft display unusual agreement patterns?*

We have seen then, that an extraposition-from-NP analysis of the *it*-cleft is supported by the historical data. On this account, the early *it*-cleft construction inherits from both the determinative pronoun construction and the OE paratactic restrictive relative clause construction. However, for many authors, the *it*-cleft's number agreement patterns pose an inexorable problem for extraposition-from-NP analyses. Nevertheless, in §4.3, I argued that while number agreement does not behave in the way we expect given the type of analysis argued for in this thesis, it can nevertheless be given an adequate explanation. In this section, I show how this particular pattern of agreement came about and provide historical evidence in support of the explanation given in §4.3.

Assuming an extraposition-from-NP analysis of *it*-clefts, the construction has an unusual pattern of number agreement. While relative clauses typically agree with their nominal antecedents, the verb embedded in the cleft clause shows number agreement with the focal NP, rather than the initial *it*. For example, in (11), the subject pronoun and

the matrix copula are marked as singular, while the focal NP (*John and Sally*) and the embedded verb (*are*) are both plural. So what is the history of this unusual number agreement pattern?

(11) It is John and Sally who are responsible

First, we need to look at the history of agreement in the matrix clause. According to Ball's (1991) study, throughout Old English and well into Middle English, the matrix copula showed agreement in number with the focal NP. For example, the excerpt in (12) contains two *it*-clefts. In the first, the form of the copula *is* agrees with the singular referent *your own spirit*. However, in the second instance, the matrix verb and the postcopular NP are both marked as plural.

(12) *...when it <u>is þin owe spirite</u> þat spekiþ þees iueles, or it <u>ben þees oþer iuel spirites</u> þat speken hem in þee.* (e15th, *A Tretis of Discrescyon of Spirites*, 88.4)
'...when it is your own spirit that speaks these evils, or it are these other evil spirits that speak them in you' (Ball 1991: 286)

By the early 15th century, we begin to get the situation we have today, with the copular verb agreeing with the subject pronoun. Example (13), which is from a slightly later manuscript of the same text, shows the more modern pattern. Here, the postcopular NP is plural but the matrix verb is in the singular form.

(13) *...or it <u>is þees oþer iuel spirites</u>*...(mid 15th, *A Tretis of Discrescyon of Spirites*)
'...or it is these other evil spirits... (Ball 1991: 286)

So how do we account for the *it*-cleft's early agreement pattern and what can it tell us about the present-day *it*-cleft? Once we relate this data to the synchronic language system of the time, the matrix-clause agreement in (12) gains a simple explanation. As

Ball (1991: 154) notes, during Old English and Early Middle English, the pronoun *hit* could be used with plural reference. For example in (14), *hit* (it) is used alongside *ha* (they) to refer to the plural referent, the *clever and wise words*.

(14) *...witti ant wise w[o]rdes hit weren ȝef ha neren false...*

(1200-25 *Seinte Katerine* (B), 114)

'...clever and wise words they would be if they weren't false' (Ball 1991: 154)

In (12) then, it is not just the focal NP that the matrix copula agrees with. Instead, number agreement is consistent across the whole configuration, including the subject pronoun *it*.

The early English data therefore shows that, at one time, the pronoun *it* was maximally underspecified; that is, it could describe or refer to plural as well as singular individuals and to human as well as non-human referents (see example (5) above). The historical flexibility of the pronoun *it* provides an explanation for some of the present-day *it*-cleft's more idiosyncratic properties. For example, although the pronoun *it* in present-day English is restricted to denoting nonhuman, singular individuals, it is well-noted that the *it*-cleft permits both human and plural foci; that is, the discontinuous description headed by the initial *it* can describe (or be predicated of) plural and/or human referents. This suggests that while the cleft *it* is now morphologically singular, it is nevertheless reminiscent of an earlier stage of the language when the pronoun *it* was maximally underspecified. Undoubtedly, it is this property of the initial *it* that enables the *it*-cleft to be so very productive. I come back to this issue in §8.2.

The diachronic data therefore supports an analysis of the present-day *it*-cleft in which the initial *it* is semantically underspecified and, although formally marked as such, is not limited to describing nonhuman and singular referents. However, the fact remains that the cleft *it* is now morphologically singular, with the agreeing matrix verb consistently occurring in the singular form. It should therefore follow that if *it* is the antecedent to the cleft clause, then the embedded verb should show agreement with *it*.

As shown in (11) above, this is not the case. So how do we explain this irregular agreement pattern?

On first appearances, the diachronic evidence does not seem overly helpful. Ball's (1991) data indicates that throughout the entire history of the *it*-cleft, the verb in the cleft clause agrees with the focal NP, rather than with the initial *it*. However, there is an interesting exception to this rule which, along with Ball's account of it, is especially relevant both to the particular *it*-cleft analysis argued for in this thesis and the explanation of number agreement given in §4.3. Ball identifies seven *it*-cleft examples with plural foci within her Late Middle English corpus. For all but one of these tokens, the verb in the cleft clause agrees with the focal NP. The exception to the rule is given here as (15). In this example, the pronoun *ye* is formally plural but the embedded verb *hath* is marked as singular.

(15) *Truly, syre, she sayd, I trowe hit be not ye that hath slayne my husband, for he that dyd that dede is sore wounded, and he is neuer lyckly to recouer, that shal I ensure hym.* (1485 Malory's *Le Morte d'Arthur* (Cx), 152.9)
'...it is not you that has slain my husband...' (Ball 1991: 307)

Ball explains this agreement pattern in the following way. She notes that while the pronoun *ye* is formally plural, in this example, it has singular reference. As a result, she concludes that "the embedded verb agrees in number with the referent of *ye*" (Ball 1991: 307).

This suggests that, rather than relying on formal, morphosyntactic agreement between the embedded verb and the focal NP, number agreement in *it*-clefts can be governed by semantic factors. This supports the semantic explanation for number agreement given in §4.3. Here, I argued that, as a collective entity, the discontinuous description in specificational *it*-clefts with plural foci has two possible conceptualizations: on the one hand, it denotes a singular set, while on the other, it denotes the plural members of this set. With this analysis in place, number agreement in specificational *it*-clefts works as follows: while the matrix copula agrees with the subject

as denoting a singular set, the verb embedded in the cleft clause agrees with the membership (singular or plural) of the described set.

In truth then, it is not really the focal NP with which the embedded verb shows agreement, but the referent(s) that make(s) up the membership of the set described by the discontinuous definite description. Unlike for noun phrases introduced by a determiner-noun combination, such as *the one* or *the ones*, the cleft *it* is underspecified for number (that is, it can describe a set with one or more members). As a result, the membership of the set is specified by the postcopular expression. It is this that gives the impression of morphosyntactic agreement between the focal NP and the embedded verb.

Therefore, from the historical evidence, it seems that agreement is dependent upon the number of referents rather than the morphosyntactic marking of either the pronoun *it* or the focal NP. From this it follows that, as a semantic phenomenon, number agreement in *it*-clefts does not present an inexorable obstacle for extraposition-from-NP accounts.[8]

---

[8] In this section, I have used historical evidence to show that number agreement does not necessarily pose a problem for the extraposition-from-NP analysis argued for in this thesis. However, at first sight, the history of person agreement in *it*-clefts seems to argue in favour of an expletive account. As I explained in §4.3, person agreement in the present-day cleft supports an extraposition-from-NP account, since, for most speakers, the focal NP is in the objective case and the embedded verb is consistently 3rd person. However, when non-3rd person foci first appeared in the LME period, these pronouns were always nominative and the verb embedded in the cleft clause could sometimes show person agreement with the clefted constituent (Ball 1991: 308-309). For example, in (i), the focal NP is nominative and the embedded verb *haue* agrees with the first person pronoun *I*.

(i) *Wherefore it is onely I that haue offenced.* (1531 Elyot Gouernour II, 140)
(Rydén 1966: 316; cited in Ball 1991: 309)

The earlier agreement pattern seems to provide evidence in favour of an expletive analysis, in which the focal NP functions as the underlying subject to the verb embedded in the cleft clause. However, there is additional evidence which goes against this hypothesis. For one thing, during this same period, the postcopular referring expression in simple copular sentences was also in the nominative case (Ball 1991: 14-15). In such sentences, it is not possible to construe the focal NP as the subject of some additional clause. Furthermore, Ball (1991: 310) notes that lack of person agreement was very common between relative clauses and their determinative pronoun heads (see also Ball 1999). She provides the following example of a reverse pseudocleft with a determinative pronoun functioning as the head of the predicative NP. Here, the embedded verb shows person agreement with the subject of the matrix clause rather than with the 3rd person determinative pronoun *he*.

(ii) *Y am he that haue synned, and Y dide wickidli...* (a1420 *Wycliffite Later Version* (A), 24.17)
'I am the one that have sinned, and I acted wickedly...' (Ball 1991: 311; Ball 1999)

7.1.4 Summary and interim conclusions

The historical data offers a new data set with which to readdress the questions left unresolved from analysing present-day *it*-cleft examples. Over the last three sections, I have shown that, once we examine the *it*-cleft in relation to the language system of earlier periods of English, many of its construction-specific structural properties are shown to have been, at one time, inherited from more general linguistic patterns. In particular, the restrictive modification of the pronoun *it*, the extraposition of the relative clause and the *it*-cleft's number agreement patterns were originally motivated by more schematic constructions including the determinative pronoun construction, the restrictive relative clause construction and the semantically underspecified pronoun *it*. While these higher-order constructions have since fallen out of productive use (such as the determinative pronoun construction) or have undergone significant changes (such as the external syntax of restrictive relatives and the restriction of *it* to non-human, singular individuals) their influence has become entrenched within the *it*-cleft schema. The historical evidence therefore provides internal motivation (via inheritance) for properties which are specific to an extraposition-from-NP account of *it*-clefts. This supports the *it*-cleft analysis argued for in this thesis and, in turn, confirms its explicatory merits.

7.2 The *it*-cleft and constructional change

In §7.1, I made use of historical evidence to identify inheritance relations which are no longer productive in the constructional hierarchy. The evidence suggests that although the structured inventory of the language has changed, the *it*-cleft's configuration has been largely unaffected. While §7.1 therefore focuses on properties that have become entrenched within the *it*-cleft schema, in this section I investigate the ways in which the (specificational) *it*-cleft construction has developed and changed over time.

As I explained in §5.3 and §5.4, the range of possible *it*-cleft foci and the ability for new information to be expressed in the relative clause are construction-specific properties which are not shared by other types of specificational copular sentence. The

---

The earlier agreement pattern in the *it*-cleft is therefore indicative of larger generalizations affecting copular constructions and determinative pronouns. As a result, this historical data does not necessarily support an expletive analysis of *it*-clefts.

question of how the *it*-cleft has acquired these idiosyncratic properties is the topic of this section. As part of her investigation into the historical development of the *it*-cleft, Ball (1991) finds that the most common type of *it*-cleft in present-day English (those with nominal focal elements and given information in the relative clause) are also the oldest, and that *it*-clefts with non-NP focal elements and *it*-clefts which permit new information in the relative clause are later developments, having their origins in the Late Middle English period.

In what follows, I present a number of problems with Ball's interpretation of this data and go on to reexamine the development if the *it*-cleft from the perspective of grammaticalization theory. I show that the gradual expansion of the *it*-cleft schema (involving the accommodation of a wider range of focal elements and new information in the cleft clause) is consistent with the unidirectional changes typical of 'grammatical constructionalization'. I undertake this diachronic investigation in §7.2.2 after providing some background information on the relatively new field of diachronic construction grammar and its integration with grammaticalization theory in §7.2.1.[9]

7.2.1 Construction grammar and grammaticalization

Grammaticalization is most commonly defined as the change through which lexical items become grammatical items and already grammatical items go on to become more grammatical. However, this atomistic view has been recently revised to accommodate the current understanding that changes to the grammaticalizing element are largely dependent upon a particular linguistic context. For example, Traugott (2003: 645) defines grammaticalization as the "process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts is assigned grammatical function, and once grammatical, is assigned increasingly grammatical, operator-like function". While this characterization of the grammaticalization process is widely accepted, it nevertheless raises some important theoretical questions. If changes to the grammaticalizing element are governed by a particular linguistic context then how has

[9] The discussion in §7.2.1 has formed the basis for parts of a chapter written for *The Oxford handbook of grammaticalization* (see Gisborne and Patten forthcoming).

this context (or 'construction') changed? What is the proper unit of study in grammaticalization? Is it the grammaticalizing element or the surrounding construction?

In response to such questions, a handful of grammaticalization theorists have redefined grammaticalization as a process affecting constructions (that is, multiword linguistic patterns) rather than individual lexical items. For example, Himmelmann (2004: 31) argues that "the unit to which grammaticization properly applies are *constructions*, not isolated lexical items" (italic original); see also Bybee (2003). Himmelmann (2004: 32) defines grammaticalization as "a process of context-expansion" whereby the construction allows a wider range of components to enter into it (called "host-class expansion") and occurs in a broader variety of larger syntactic and/or semantic and pragmatic contexts. However, the notion of the grammaticalizing lexical item is still important to this definition; Himmelmann (2004: 32-33) notes that "grammaticization applies only to the context expansion of constructions which include at least one grammaticizing element".

Despite appearances then, Himmelmann's (2004) definition of grammaticalization is still centred on an atomic grammaticalizing element, but brings the surrounding context into view as well. Himmelmann considers the grammaticalization of the individual element to be dependent upon the schematicity (or generality) of the surrounding construction. As a result of context-expansion, the fixed grammaticalizing element occurs alongside "a growing class of items which enter into this construction" and consequently becomes "the increasingly general construction marker" (Himmelmann 2004: 38). From this perspective, traditional diagnostics of grammaticalization, including semantic bleaching, phonological erosion and so on, are treated as epiphenomena.

Such issues in grammaticalization theory are especially topical when considered in relation to construction grammar. Within this framework, the construction (defined here as a form-meaning pair) is understood to be the basic (and in fact, the only) unit of linguistic knowledge. From this, it follows that the framework calls for a construction-based model of language change. Relevant questions to the field of diachronic construction grammar include: how do new constructions emerge? How do existing

constructions change? And how does the organization of the constructional taxonomy change? In what follows, I examine how a construction-based model interrelates with the preexisting accounts of language change developed within grammaticalization theory.

With the rising interest in diachronic construction grammar "there have been more and more attempts to explore the relationship between patterns of constructional change, and existing accounts of grammaticalization" (Trousdale 2008c). Although grammaticalization theory traditionally focuses on changes to atomic lexical items, in construction grammar, lexical items are also classed as constructions; that is, they are symbolic units of form and meaning. From this, it follows that changes which apply to substantive, simple constructions should also affect more schematic and complex constructions. If lexical items can grammaticalize, larger, less substantive constructions should also be subject to grammaticalization, not only as the surrounding context for a grammaticalizing morpheme, but as the actual grammaticalizing element (Trousdale 2008b: 33-34).

However, if we accept a constructional model of language structure, then our definition of grammaticalization must change. For one thing, construction grammar does not recognize a distinction between the lexicon and the grammar. Within this framework, all units of grammatical knowledge, including lexical items, are given a unified representation as constructions and are located within the network that represents a speaker's grammar. This begs the question then, given the organization of the constructional taxonomy, what does it mean to say that a construction becomes more lexical or more grammatical? And how can we represent these directional changes?

According to Traugott (2007), Trousdale (2008b) and Fried (2008), the hierarchical network of constructions provides a useful means of identifying and accounting for directional changes. Within this network, specific linguistic patterns inherit properties from more basic constructions. The grammaticalization of constructions (or 'grammatical constructionalization') originates in language use with actually occurring utterances. As new tokens emerge, the speaker generalizes over these instances (or constructs) to create a new level of abstraction. This in turn can have

repercussions higher up in the taxonomy as existing schemas become more abstract in order to accommodate (or sanction) these new lower-level constructions (Trousdale 2008b: 55).

Within the framework of construction grammar, grammaticalization is therefore a process of schematization, in which the construction becomes a more abstract, higher-level category and its internal composition becomes less fixed. On this account, the cline from lexical to grammatical status is re-envisaged as a hierarchy from more substantive to more schematic constructions. As Trousdale (2008a: 170-171) comments, "The more schematic the construction, the more productive it will be (thus such constructions become aligned with what is usually called 'syntax' and 'productive morphology'); the more substantive the construction, the less productive it will be (i.e. it will become more associated with the 'unproductive morphology' and the 'lexicon')".[10]

The construction grammar model of language structure therefore offers a new way of thinking about language change. However, in many ways, this new approach shares many of the same assumptions as, and incorporates many of the findings of, grammaticalization theory. For instance, a constructional approach to language structure assumes a usage-based model of language change (see chapter 2). In this model, language change originates in language use and proceeds in incremental stages, following a directional pathway. As noted above, grammatical constructionalization is initiated in the actual utterances that speaker's produce, rather than in the abstractions developed within the speaker's mind. The change then proceeds upwards throughout the constructional hierarchy, as speakers inductively generalize over instances to form schemas, leading to the creation of new constructions and the reconfiguration of existing

[10] In addition to becoming increasingly schematic and productive, Traugott (2007) and Trousdale (2008a, b) suggest that as a construction grammaticalizes, it also becomes less compositional. For these authors, 'compositionality' refers to the extent to which the construction's meaning is predictable from the meanings associated with its component parts. However, as I explained in §2.5 and §6.2.4, an important advantage to the *it*-cleft analysis argued for in this thesis is that it is 'compositional' in the sense that the construction's meaning can be broken down and attributed to its component parts. As a result, in this thesis, I prefer to talk about the outcome of grammatical constructionalization as involving the development of construction-specific meaning and mismatched correspondences, rather than a loss of compositionality. This practice actually allows us to be more precise about the exact nature of constructional change in the *it*-cleft's development.

ones. In this way, constructional change is both gradual (proceeding in incremental steps) and directional.

Construction grammar as a theoretical framework is therefore able to intersect with and work alongside theories of language change. This represents an important advantage over parameter-based theories of grammar, such as minimalism. For example, while Roberts and Roussou (2003) attempt to accommodate the findings of grammaticalization theory into a minimalist account of syntactic change, they nevertheless acknowledge that an account based upon parameter setting is difficult to reconcile with the empirical evidence for pathways of change.

The Principles and Parameters theory of grammar assumes that language change happens across generations, during the acquisition process. Once the speaker has acquired their language (through parameter setting) their core grammar is fixed and is not susceptible to change. However, a child may acquire a different grammar from their parents depending upon their linguistic input. The parameter setting account therefore assumes that language change is both abrupt and random (see Lightfoot 1979, 1999). However, the diachronic evidence suggests otherwise. As Roberts and Roussou (2003: 4) comment, "the phenomena of grammaticalization provide ample evidence of just such pathways or tendencies, and thus syntactic change must – at the very least at the descriptive level – be in a significant sense non-random."

The basic assumptions of the P&P framework are therefore at odds with the findings of grammaticalization theory. As a result, while Roberts and Roussou (2004) attempt to reconcile their model of language change with the empirical evidence, proposing a theory of markedness which creates "basins of attraction" within the parameter space, their effort focuses on how they can make grammaticalization fit into minimalist theory rather than how their theory of grammar might be of use to those working on grammaticalization. The construction grammar framework, on the other hand, offers a new perspective from which to view grammaticalization changes while retaining the fundamental principles upon which the diachronic theory is based. For instance, according to Trousdale (2008c), a constructional approach allows us to capture

generalizations which are not apparent from the traditional atomistic view of the grammaticalization process. He provides the following example of the English modals.

Viewed from the perspective of the atomic lexical items, the English modals are the product of individual changes which conform to the well-established grammaticalization cline: lexical verb > modal verb. However, Warner (1993) claims that this development was far from straightforward. He finds that, during the transition from Old to Middle English, the modals actually developed more verb-like properties (such as the development of new present tense forms). As Trousdale (2008c) comments, Warner's (1993) findings therefore seem to contradict the unidirectionality hypothesis.

In contrast, a constructional approach treats the grammaticalization of each verb as the creation of a new construction. As more and more verbs develop modal uses, a new overarching abstraction (the modal construction) emerges. As the modal category expands, sanctioning new subtypes, this in turn has implications for the higher-order auxiliary construction, which becomes a more basic and distinctive category (see Warner 1993 and Hudson 1997 on the "strengthening" of the auxiliary and modal categories). On this account, unidirectionality (which is understood here as relating to the creation and reconfiguration of constructions within a network) is preserved; although individual items may have acquired more verb-like characteristics, the modal and auxiliary constructions undoubtedly became more schematic and productive over time. As Trousdale (2008c) notes, while "the standard conceptualization of the cline" is itself a generalization, it focuses on the development of the atomic lexical item, and so "fails to consider the larger constructional changes within which such micro-changes are embedded".

An additional advantage to the constructional approach is that it allows us to be more precise about what we mean by the terms "more lexical" or "more grammatical". Grammaticalization theorists often find it difficult to draw a sharp line between items that are properly in the lexicon and items that are located in the grammar (Himmelmann 2004: 25). Indeed, the very concept of grammaticalization as a change which makes lexical items more grammatical argues against a theory of language structure which rigidly separates the lexicon from the grammar. In construction grammar, we have the

natural extension of this idea: grammatical constructions (as units of form and meaning) have the same status as lexical items, both of which are incorporated into the hierarchical network which makes up a speaker's knowledge of their language. Within the hierarchical taxonomy of constructions, "more lexical" and "more grammatical" are measured in degrees of productivity and schematicity. This more accurately corresponds with the notion that change is gradual and directional, progressing in incremental stages along a continuum.

Therefore, once we examine grammaticalization changes in relation to the constructional taxonomy, the process is reinterpreted as 'schematization' or 'expansion'. In this way, the constructional model of change overlaps considerably with the accounts of grammaticalization provided by Bybee (2003) and Himmelmann (2004). The question of whether grammaticalization theory should be entirely subsumed by a constructional model of change or whether the term 'grammaticalization' should be reserved for changes affecting only atomic elements remains an open issue.[11] Nevertheless, what is clear is that there is a relationship (of some sort) between constructional changes and traditional grammaticalization phenomena.

I return to this issue in chapter 8. In the next section, §7.2.2, I show that the gradual emergence of new subtypes of *it*-cleft conforms to the expected pattern for grammatical constructionalization. In §8.1, I identify well-known diagnostics for grammaticalization in these changes and ask what the diachronic development of the *it*-cleft can tell us about the nature of constructional change and its relationship to grammaticalization.

7.2.2 A diachronic investigation of non-NP it-*clefts and IP* it-*clefts*

As I noted above, Ball (1991, 1994a) finds that *it*-clefts with non-NP foci and *it*-clefts with new information in the cleft clause have their origins in the Late Middle English period. Ball speculates that this development involved a series of mergers between

[11] For example, Noël (2007) argues that while the processes of schematization (or constructionalization) and grammaticalization often intersect, they should nevertheless be treated as separate developments. He claims that, while the schematization of larger constructions often leads to the grammaticalization of simple atomic elements, these are two different types of change.

multiple configurations including the existing NP-focus *it*-cleft, an Old English impersonal construction and the reverse pseudocleft. In this section, I identify a number of problems with Ball's interpretation of the diachronic data. I show that Ball's expletive account of the present-day *it*-cleft leads her to select an unnecessarily complex and implausible analysis over an intuitively simple one. Using data from the Penn Parsed Corpora of Historical English, I reexamine the development of the *it*-cleft as a straightforward example of grammatical constructionalization. Non-NP *it*-clefts and informative-presupposition (IP) *it*-clefts are shown to originate by extension from existing *it*-cleft constructs, overriding more general patterns of correspondence. The upshot is that the overarching specificational *it*-cleft construction becomes increasingly schematic and productive over time.

The section is structured as follows. In §7.2.2.1, I present and criticize Ball's account of the development of *it*-clefts with non-NP foci before providing my own evidence for a grammatical constructionalization story in §7.2.2.2. In §7.2.2.3, I discuss Ball's account of the development of the informative-presupposition *it*-cleft before outlining my own analysis of the corpus data in §7.2.2.4.[12]

7.2.2.1 Ball (1991, 1994a) and the AdvP/PP *it*-cleft

Ball (1991, 1994a) treats *it*-clefts with prepositional phrase and adverb phrase foci together, as making up a construction (the AdvP/PP *it*-cleft) which is separate from the NP-focus *it*-cleft. She suggests that this new kind of *it*-cleft evolved from an Old English sentence-type with *beon/wesan*, which is traditionally classified as an impersonal construction. Ball notes that these impersonals could sometimes occur with adjuncts; for instance, in (16), the string *not long afterward* is not an obligatory component of the sentence. Such examples show a superficial similarity to modern day *it*-clefts with adverb phrase and prepositional phrase focal elements, which often express relations such as time or place, as in (17).

[12] Large parts of the material in §7.2.2 are to appear in Patten (in press).

(16) Was hyt nat long aftyrward,
He fyll yn a sykenes hard. (c1400, Mannyng, *HS(B)*, 5459)
'Not long after that, he fell into a grave sickness' (Ball 1994a: 612)

(17) A: How long has he been feeling this way?
B: Well, he lost his job six months ago and it was shortly after that that we noticed a distinct change in his temperament.

However, despite their apparent similarity, these impersonal examples are not *it*-clefts. In the impersonal construction, the verb is not the copular *be* that we find in cleft sentences; instead it has a full meaning, which can be glossed as something like 'happen' or 'come to pass' (Ball 1994a: 611). So, for example, the meaning of Ball's example presented here as (16), is something like 'Not long after, *it happened that* he fell into a grave sickness'. That Late Middle English examples such as (16) are impersonals and not clefts is supported by the fact that they have variants with other *happen*-class verbs, such as *fallen*. For example, (18) can be glossed as 'At the end of twelve months, *it befell that* his companions went to the mountain'.

(18) Fyl hyt at þe twelue moneþ ende,
Hys felaus to þe mounteyne gun wende, (c1400, Mannyng, *HS(B)*, 10765)
'At the end of twelve months, his companions went to the mountain'
(Ball 1994a: 612)

Ball recognizes that these impersonals do not require a cleft analysis, but suggests that during the Late Middle English period they underwent a partial merger with the existing NP-focus *it*-cleft, resulting in a new and separate construction, the AdvP/PP *it*-cleft. At this point, we begin to get examples with non-NP focal elements which are clearly clefts rather than impersonals, such as (19).

(19) Of hym is all, for he is creatour,
Be hym it is þat all þing is susteyned,
In hym is all þing kyndely conteyned. (e15th, Walton tr., *Bo*, 4.preface.2.5)
'Everything is from him, for he is creator; it is by him that all things are sustained; all things are naturally contained in him.'
(Ball 1994a: 613; my translation)

In support of a cleft analysis for this example, Ball (1994a: 613) notes that the prepositional phrase *by him* is not quite as optional as the adjuncts found in impersonals. She argues that if we remove this element, the remaining sentence in (20) does not really express a well formed proposition. The reason for this is that in (19), the prepositional phrase focal element is much more important to the understanding of the sentence, since it expresses agency rather than time or location.

(20) ??All things are sustained (Ball 1994a: 613)

Ball's (1991, 1994a) explanation of how the AdvP/PP *it*-cleft originated is not unreasonable, since, as we have seen, the impersonal construction shows a strong superficial similarity to the *it*-cleft. However, a much more straightforward hypothesis is that *it*-clefts with non-nominal foci simply represent an extension of the existing NP-focus *it*-cleft and were not directly influenced by the Late Middle English impersonals. Nevertheless, as a result of her particular assumptions about the present day *it*-cleft construction, Ball does not even consider this as a possibility.

Although Ball's (1991, 1994a) focus is exclusively diachronic, she nevertheless subscribes to a particular synchronic account of the present-day *it*-cleft. As I explained in §7.1.1, Ball assumes an expletive analysis, in which the focal element is interpreted in relation to the gap in the cleft clause and the constituents *it* and *be* are semantically empty. For Ball (1994a: 605) then, *it*-clefts represent information-structure variants of, but are nevertheless equivalent in meaning to, their corresponding noncopular counterparts.

(21) It was [[the therapist]$_i$ [that ____$_i$ killed her]] [NP-focus *it*-cleft]

(22) The therapist killed her [canonical counterpart]

Ball analyses the cleft clause in NP-focus *it*-clefts as a restrictive relative.[13] However, for *it*-clefts with non-nominal foci, Ball (1994a: 605) claims that "there is little support for a restrictive relative analysis of the subordinate clause". Ball follows Delin (1989) in suggesting that these examples contain a "sentential complement" rather than a relative clause. This analysis is supported by the fact that the *that*-clause in non-NP *it*-clefts often expresses a complete sentence, without a perceptible gap, as in (23). From this, it follows that the present day *it*-cleft is not a unified construction for Ball and that *it*-clefts with non-nominal focal elements make up a separate construction from the NP-focus *it*-cleft.

(23) It's in December that she's coming [PP-focus *it*-cleft]

By adopting this synchronic perspective, Ball is forced to find a historical explanation for examples with non-NP focal elements that goes beyond the simple extension of a single *it*-cleft construction. By incorporating impersonals into her analysis, Ball can argue that adverb phrase and prepositional phrase focus *it*-clefts were, from the very beginning, a separate construction. Since impersonals have the function of presenting a complete sentence, with or without an optional adjunct, this can explain why the *that*-clause of non-NP focus *it*-clefts should be analysed as a sentential complement rather than as a relative clause.

Ball's historical explanation is therefore in some way dependent upon, or is at least motivated by, a particular synchronic analysis. However, a more important criticism is that Ball's account of the development of the AdvP/PP *it*-cleft also suffers

[13] However, as Ball (1994a: 640f) observes, "A relative clause analysis for the complement of the NP-focus *it*-cleft is not uncontroversial" on an expletive account, since the focal NP in *it*-clefts can be a proper noun or pronoun, elements which cannot normally be modified by a restrictive relative clause (see §4.1.3). For this reason, Ball (1977, 1991) claims that while the cleft clause is structured internally like a restrictive relative, it is a "headless relative" which does not modify its immediate antecedent (see §5.2.1 and §7.1.1).

from a lack of plausibility. A problem with Ball's story is that it is not immediately clear why the impersonal construction and the Late Middle English *it*-cleft would merge in the first place. It is true that these impersonals show a superficial similarity to modern day adverbial and prepositional clefts. However, before this particular type of *it*-cleft existed, it is difficult to see what the two constructions share.

The impersonal construction and the NP-focus *it*-cleft are very different, both in terms of structure and meaning. In the superficial relationship that Ball exploits, the focal noun phrase of the *it*-cleft corresponds to the adjunct in the impersonal examples. However, these adjuncts are made up of non-nominal word classes. Furthermore, they are purely optional, are not focal elements, and indeed have very little bearing on the sentence. As indicated above, although both constructions involve forms of *be*, that found in impersonals is not the copula of cleft sentences, but is a verb of full meaning. Likewise, the sentence-final relative clause of the NP-focus *it*-cleft, containing a gap, is structurally distinct from the syntactically complete sentential component present in the impersonal construction.

These elements also differ in information status. The relative clause of all types of *it*-cleft is presupposed and, for NP *it*-clefts in the pre-modern period in particular, is largely restricted to expressing given information. In contrast, the sentential component in impersonals is not presupposed and in fact makes up the main informational content of the construction. These two sentence types therefore differ in function: while clefts are well known for packaging information in a marked way, impersonals represent a neutral way of presenting information. Consequently, for Ball, the only similarity that NP-focus *it*-clefts share in terms of function with impersonals, is an expletive *it*. However, this property of clefts is peculiar to the particular present day analysis that Ball subscribes to. Since I have argued against an expletive analysis for *it*-clefts throughout this thesis, for me, *it*-clefts and impersonals have even less in common.

7.2.2.2 A constructionalization story for the development of non-NP *it*-clefts

In contrast to Ball, I do not assume that modern day *it*-clefts with non-nominal foci make up a separate construction from those with noun phrases in the postcopular position. Rather than relating *it*-clefts to simple noncopular sentences, I assume that *it*-clefts correspond closely to other specificational copular constructions. This has led me to argue in favour of an extraposition-from-NP analysis of *it*-clefts. As I explained in §4.1.3 and §5.3, once we analyse the initial *it* as the antecedent to the cleft clause, there is no problem with incorporating a relative clause structure into clefts with focal elements that resist restrictive modification in other constructions. Regardless of whether there is a perceptible gap in the relative clause or not, all types of *it*-cleft can be analysed in the same way. For instance, the prepositional phrase focus *it*-cleft in (24) is interpreted as containing a definite-like description, akin to that in the specificational *NP be NP* sentence given in (25).[14] From this, it follows that all *it*-clefts, regardless of focus category, form part of a single construction.

(24) [It] is in December [that she's coming] [PP-focus *it*-cleft]

(25) [The time/date [that she's coming]] is in December [*NP be NP* sentence]

This unified account of the present-day *it*-cleft supports a historical analysis whereby non-nominal focal elements were gradually accommodated into the *it*-cleft construction via coercion. I provide evidence for this grammatical constructionalization story using data from the Late Middle English texts in the Penn-Helsinki Parsed Corpus of Middle English second edition (PPCME2) and from the complete Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME).[15] Frequency counts of *it*-clefts with different categories of focal element are given in Table 7.1. The data is separated into time periods recognized by the Penn-Helsinki Parsed Corpus which rely on composition

[14] The example in (25) has both a specificational reading, such that *the time* is/equals *in December*, and a predicational reading, such that *the time* is located *in December*. The specificational reading is what is relevant here.

[15] The data was retrieved by searching for the syntactically annotated *it*-cleft structure using the CorpusSearch 2 program. See §1.5 for information regarding the size and makeup of the corpora.

dates (where known).[16] This table indicates both raw numbers and frequencies normalized to a corpus of 500000 words (shown in parentheses). The numbers given here are quite conservative; I have only included clear cases of specificational *it*-clefts. For example, predicational clefts, existential sentences and instances which prefer an impersonal reading are not included in these frequency counts.[17] These examples require a separate analysis.[18]

---

[16] The PPCEME is separated into 3 time periods (E1, E2 and E3), which are represented in Table 7.1. For the PPCME2, the corpus is divided into subcorpora indicating both composition and manuscript date. In Table 7.1, I conflate these subcorpora to indicate only the time of composition. However, while for the PPCME2 directory, the period M2 ranges from 1250-1350, I only include texts during the LME period from 1300-1350.

[17] Existential sentences at this time could occur with *it* as well as *there*. For instance, the rhetorical question below asks *what is there* (or *what exists*) *that a man can do to another man which cannot be done, in turn, to him.* The expected answer to this question is of course *nothing*. The follow-up sentence gives an example of this generalization by reminding us that Busiris, who frequently killed his guests, was in turn killed by one of his guests – Hercules.

(i) To go further, what is it that any many may do to another, which another may not do again to him? We are told that it was the Custom of Busiris to kill his Guests, and himself at last was killed by Hercules his Guest. (Preston's Boetius (Boethpr), 1965)

While such sentences are superficially similar to specificational *it*-clefts, they are pragmatically distinct. Existential sentences question rather than presuppose the existence of the entity described in the *that*-clause (Ball 1991: 269). Consequently, I have excluded these tokens from my specificational *it*-cleft data-set.

[18] In addition, I have omitted eleven tokens introduced by the pattern *I it am* from the PPCME2 search results. Ten of these, shown below, are from a single passage of the same text.

(i) …ofte tymes oure lorde Ihesu sayde to me, "I it am that is hiaste. I it am that þou luffes. I it am that thowe lykes. I it am that þowe serves. I it am þat þou langes. I it am that þowe desyres. I it am that thowe menes. I it am þat is alle. I it am that haly kyrke preches the and teches the. I it am that schewed me are to the". (Julian of Norwich's Revelations of Divine Love, c1450)
'…often times our lord Jesus said to me, "I am the one that is the highest. I am the one that you love. I am the one that you like. I am the one that you serve. I am the one that you long (for). I am the one that you desire. I am the one to whom your attention is directed. I am the one that is all. I am the one that holy church preaches to you and teaches you. I am the one that showed myself to you before."'

Although these tokens look like focus-first specificational *it*-clefts, they are instead examples of a predicate nominal construction. The function of these tokens is to tell us more information about the subject referent (*Jesus*), rather than to identify the referent described by the definite-like description. Here, the initial pronoun *I* is the subject, which shows agreement with the copular verb *am*. The sentence-final relative clause modifies the pronoun *it* to form a definite NP predicate. According to Ball (1991: 71) the configuration *I it am* is therefore equivalent to *I am he*. As Watson and Jenkins (2006: 206) comment, in the example above, the use of *I it am* rather than *I am he* may be a way of deemphasizing the maleness of Jesus at a time when the gender-neutral construction *I am the one* had not yet developed.

Table 7.1 *The frequency of* it-clefts *with a range of foci throughout LME and EME*

	NP	PP	AdvP	CL *(because)*	Total
M2 (1300-1350)	3 (10.5)				3 (10.5)
M3 (1350-1420)	11 (11.3)	1 (1.0)			12 (12.3)
M4 (1420-1500)	16 (30.2)				16 (30.2)
E1 (1500-1569)	24 (20.8)	1 (0.9)		1 (0.9)	26 (22.6)
E2 (1570-1639)	37 (28.3)	2 (1.5)			39 (29.8)
E3 (1640-1710)	73 (64.6)	18 (15.9)	8 (7.1)	1 (0.9)	100 (88.5)

There are several characteristics of this data that suggest that non-NP *it*-clefts are extensions of the existing NP-focus *it*-cleft rather than the result of a merger between the *it*-cleft and the impersonal construction. First, the earliest examples of *it*-clefts with prepositional focal elements do not contain temporal or spatial adjuncts, but express relations such as agency, as in Ball's example given in (19) above, or cause, such as the PPCME2 example in (26). Tellingly, these relations can also be expressed by NP-focus *it*-clefts during the Late Middle English period, shown in example (27).

(26) *Me troweþ þat <u>by þe prayers of þis holy mayde</u> it is þat þat place was never*
Me believes that by the prayers of this holy maid it is that that place was never
ȝit detstroyed. (John of Trevisa's Polychronicon, a1387)
yet destroyed
'I think that it was by the prayers of this holy maiden that that place was never destroyed'

(27) *It es pride in þaim þat hyes þaim.* (Rule of St Benet, a1425)
It is pride in them that highs them
'It is pride that elevates them'

Like (19), (26) cannot be given an impersonal interpretation because the phrasal element (underlined) contributes the most important information; it cannot be removed from the

sentence and still leave the essential meaning intact. For instance, in (26), the writer expresses his thoughts that Ethelberga's prayers prevented the destruction of the abbey. If (26) was interpreted as an impersonal, with an optional adjunct, we would be left with something like (28), which expresses a doubt about whether it is a fact that the abbey was never destroyed. This is a far less intuitive gloss of (26), which seems instead to presuppose the fact that the abbey was never destroyed.

(28) I think it happened that that place was never destroyed.

This suggests that examples such as (19) and (26) are more likely to have developed as extensions of the NP-focus *it*-cleft and were not directly influenced by the impersonal construction.

Examples which contain temporal and spatial foci and which are clearly non-NP *it*-clefts, despite allowing an impersonal reading, do not appear until the second half of the Early Modern period. In such cases, exemplified by (29), it is the context that reveals that they are clearly presuppositional clefts. This prevents an impersonal reading for (29), such as the possible paraphrase in (30). If impersonals really were the driving force behind the appearance of non-NP *it*-clefts, we would expect examples such as (29), which are ambiguous in form, but disambiguated by context, to be the very earliest to occur. However, as indicated above, this is not the case.

(29) Lee. and Mr. Ireland upon Saturday Night came to my Master's House…
L.C.J. What time and what Year was it?
Lee. It was in 78; but I am not certain of the Day of the Month: It was on a Saturday he came (Oates, 1685)

(30) It happened, on a Saturday, that he came to my master's house

Furthermore, we find NP-focus *it*-clefts expressing temporal and spatial relationships during this same time period, shown in (31). This suggests that NP-focus *it*-clefts and non-NP *it*-clefts are instances of a single construction.

(31) It was, as near as I can remember, the 3d of August, that he went out of Town. (Oates, 1685)

I argue that by allowing a wider range of elements into what was previously a strictly nominal position, the *it*-cleft becomes a more schematic construction. In other words, the construction loses a syntactic constraint, with the focal position no longer specifying for an NP. Nevertheless, the semantic condition for the postcopular element is still in force; that is, only elements which can perform a referential function will be allowed into this position (see §5.3). Since some syntactic categories are better suited to performing a referring function, and furthermore, some members of these categories fulfill this role better than others, this explains why changes to the focus category in *it*-clefts are gradual, occurring in incremental stages.

Tellingly, in both Ball's data and my own, the only clear cases of non-nominal *it*-cleft in the Late Middle English period involve prepositional phrases. As shown in Table 7.1 above, the adverb phrase focus *it*-cleft seems to be a later development and can only be placed, with empirical support, as originating at the end of the Early Modern era. The fact that prepositional phrases are amongst the earliest to be permitted into this position is unsurprising, since as Borkin (1984: 136) comments, there is a semantic overlap between the phrasal categories of NP and PP; although prepositional phrases typically express relations, such as location and quality, they nonetheless relate to nominal concepts such as time, place and manner.

For example, when occurring as the focal element of an *it*-cleft, Bolinger (1972) notes that prepositional phrases lose their relational properties. In (32), the meaning is not that *the time* is located *in the winter*, but that *the time* is/equals '*in the winter*'. In this example then, the relational prepositional phrase is coerced into performing a referential

function.[19] Consequently, these kinds of *it*-cleft have little difference in meaning from those with NP-focus, as in the example given in (33).

(32) If it bee in the winter that your Hawke batheth, when no sunne shineth, you may then drie her as well by the gentle aire of the fire as otherwise. (Markham, 1615)
'If it is in the winter that your hawk bathes, when no sun shines, you may then dry her as well by the gentle air of the fire as otherwise'

(33) and if this be done in November, it will preserve the Trees for that whole year, with that once doing, it being the winter time only that they will feed upon the bark. (Langford, 1696)

An important advantage to this diachronic account is that it is able to explain the continuing development of the construction, as it goes on to accommodate a much wider range of focal elements. In contrast, Ball's analysis is very much restricted to the Late Middle English period, and she has to invoke an increasing number of outside pressures to speculate about the subsequent history of the AdvP/PP-focus *it*-cleft. For example, Ball (1994a: 614) notes that, "From the LME period onwards, the *it*-cleft construction has taken in a greater variety of non-NP foci, possibly in response to the decline of some alternatives and functional change in others (e.g. preposing)".

A grammatical constructionalization story for the *it*-cleft is also compatible with, and explains, ongoing coercion effects in present-day English. For instance, as I explained in §5.3, É. Kiss (1998: 262) notes that phrasal categories which are less suited to performing a referring function can be "individualized" and made into more discrete entities via listing. This allows the adjective *sick* to be accommodated into the focal position of (34), creating a much more acceptable *it*-cleft than its counterpart in (35).

---

[19] Note too that some prepositional phrases can have a referential function in other constructions, such as the passive sentence given below. Again, this helps to explain why prepositional phrases were the earliest non-nominal category to occur in the focal position and why they are still the most frequent non-nominal category among present-day *it*-clefts.

(i) Under the stairs was painted first

(34) It's not sick that he was but tired (É. Kiss 1998: 262)

(35) *It's sick that he was (É. Kiss 1998: 262)

Likewise, Borkin (1984: 127) comments that focusing adverbs, such as *just* and *alone* can be used to "underscore the correctness of the identification" and "make explicit the reason for focusing". This explains why the inclusion of *just* improves the acceptability of the clausal-focus cleft in (36).

(36) It's [just] that he's so annoying that bothers me

Analysing the gradual emergence of non-NP *it*-clefts as a case of "host-class expansion" (see §7.2.1) is further supported by the historical development of *it*-clefts with nominal foci. In the Late Middle English data of the PPCME2, 24 out of the 30 examples of NP-focus *it*-clefts involve proper names and pronouns. This suggests that the *it*-cleft originally showed a preference for focusing the most referential and discrete of entities.[20] When more abstract nouns, which do not denote discrete physical objects, begin to appear with more frequency in the Early Modern period, we find that listing and focusing adverbs are used to accommodate these foci in the same way as they are in modern day clausal and adjectival clefts. For instance, 7 out of the 22 examples of Early Modern *it*-clefts focusing abstract nouns involve listing, such as (37), while a further 8 examples, including (38) below, contain focusing adverbs.

(37) ...that it was not the pure knowledg of nature and vniuersality...which gave the occasion to the fall; but it was the proude knowledge of good and euill...which was the fourme of the temptation; (Bacon, 1605)

(38) 'Tis use alone hardens it and makes it more able to endure the cold, (Locke, 1685)

[20] Ball (1991: 272) finds this same preference in her own LME data. She notes that 89 out of 111 NP-focus *it*-clefts contain animate foci.

From this evidence, we can therefore conclude that by allowing a wider range of elements into the focal position, the *it*-cleft construction undergoes gradual, and perhaps continual, grammaticalization (or grammatical constructionalization), showing increased schematicity. The construction loses a syntactic constraint, with the focal position losing its NP specification and becoming an open slot which any phrasal category (XP) can fill. However, the *it*-cleft retains the semantic condition that the focal element must be referential and it is this property which governs the course and rate of the *it*-cleft's development. In line with Michaelis' (2003, 2004) work on mismatch phenomena (introduced in §2.4), we might say that as non-NP and less discrete elements are accommodated into this referential slot via coercion, their nominal characteristics are emphasized and relational or situational properties are lost.

7.2.2.3 Ball (1991, 1994a) and the IP *it*-cleft

According to Ball (1991, 1994a), the development of the informative-presupposition *it*-cleft, in which new information is presented in the relative clause, is also influenced by the Late Middle English impersonal construction. Ball assumes that since the sentential complement of impersonals contains new information, this property is carried into the new AdvP/PP *it*-cleft. This construction then merges with the existing NP-focus *it*-cleft, resulting in another new and separate construction: the NP-focus IP *it*-cleft. However, Ball recognizes an important problem with this analysis: the frequency of the new AdvP/PP *it*-cleft is way too small during the LME period to influence the already established NP-focus *it*-cleft. As a result, Ball concludes that other constructions must also have played a role in this development, likely candidates being the impersonals again as well as perhaps the reverse pseudocleft.[21]

Ball's analysis of this development seems to be greatly influenced by Prince's (1978) article, which identifies two separate types of modern day *it*-cleft: the stressed-

[21] Ball (1991, 1994a) uses the term 'pseudocleft' to cover *wh*-clefts as well as *NP be NP* sentences introduced by *the one that*... In Late Middle English, the language had not yet developed these constructions. Nevertheless, Ball uses this same term to refer to their precursors, which include specificational copular sentences introduced by restrictively modified determinative pronouns, such as *he that*... 'Reverse pseudoclefts' show the opposite alignment, with the restrictively modified NP occurring in postcopular position.

focus *it*-cleft, with given information in the relative clause, and the often overlooked informative-presupposition *it*-cleft. Ball (1994a) emphasizes the functional contrast between these two types of cleft and consequently suggests that the IP *it*-cleft originated as a distinct construction from the SF *it*-cleft. That is, she assumes that the IP *it*-cleft is the result of multiple constructions conspiring together and is not a simple extension of the existing *it*-cleft construction. Tellingly, Ball acknowledges that her analysis is somewhat *ad hoc*. She notes that, "It is unlikely that we shall ever be able to pinpoint the cause of the appearance of the IP NP-focus *it*-cleft, but the two constructions together [the cleft/impersonals and the reverse pseudocleft] will have provided a sufficient condition for this development" (Ball 1994a: 621).

However, the main problem with Ball's theory is that it cannot explain the "known-fact" effect. Even when the information in the relative clause of *it*-clefts is not known to the intended audience, there is still a sense that this information is presupposed and is assumed to be true. For example, in (39) below, the audience is not aware that *someone once said 'laws are silent at times of war'*, but they are nevertheless expected to accept this information as a fact. As Prince (1978: 899-900) comments, these IP *it*-clefts "MARK A PIECE OF INFORMATION AS FACT, known to some people although not yet known to the intended hearer" (emphasis original).

(39) (Start of lecture)
It was Cicero who once said, 'Laws are silent at times of war'. [IP *it*-cleft]

If impersonals really were the driving force behind this change, it is difficult to see why the information in the relative clause would be marked in this way. In §7.2.2.1, I presented a number of arguments that the impersonal construction did not directly influence the development of the AdvP/PP *it*-cleft. In particular, unlike the relative clause of *it*-clefts, the sentential complement of the impersonal construction is not presuppositional. Consequently, it is unlikely that impersonals could have played an important role in enabling the *it*-cleft to occur with new, but presupposed information in the relative clause.

In contrast, Ball's speculation that the reverse pseudocleft may have influenced the development of the IP *it*-cleft is much more plausible, since these are both specificational copular constructions containing a presupposed relative clause. However, as Ball (1994a: 618) comments, it is difficult to find a "motivation in LME for a transfer of functions, since the cleft and the pseudo-cleft co-existed for hundreds of years before the first IP clefts appear". Furthermore, on closer inspection, Ball's examples of reverse pseudoclefts containing discourse-new information may not even be specificational. For instance, Ball (1994a: 616) notes that "Where the subject NP is not contrastive, the inverted pseudo-cleft has long been used to present information that is not Old or Inferrable from the context". However, if the subject NP is not contrastive and presents an individual that is already the topic of conversation, it is difficult to separate these examples from ordinary subject-predicate (topic-comment) sentences, as shown in (40).

(40) (a) This is John. He is a doctor.
(b) This is John. He is my best friend.
(c) This is John. He is the person/one that mends my shoes.

For copular sentences with a definite noun phrase predicate, the example will only be interpreted as having a specificational (or identifying) meaning when the subject is marked as the focus, as in (41).

(41) (a) A: Who is your best friend? B: JOHN is my best friend.
(b) A: Who stole the money? B: JOHN was the one that took it.

In the absence of prosodic clues, Ball's comment that the subject NP of her examples is not contrastive suggests that this element may be the topic rather than the focus of the sentence. For instance, Ball provides the following Old English example, in which the propositions expressed by the relative clauses are not given by the previous discourse but are nonetheless known information.

(42) Hlyst nu placida. Ic eom hælende crist…and ic eom se ðe gesette dagas. and tida. and gear. and ic eom se ðe man of eorðan gehiwode. and for manncynnes hælo. ic com to eorðan… (*LS* 8, St. Eustace, 59)

'Listen now, Placidas. I am Jesus Christ…and I am the one who set days and seasons and years. And I am the one who formed man out of the earth, and for mankind's salvation I came to earth…' (Ball 1994a: 617)

As Ball (1994a: 616) comments, in this example "Christ appears to Placidas and uses inverted pseudo-clefts to identify himself". However, identifying, or providing more information about, a topical referent is not a specificational function. In specificational sentences we have the reverse situation, with the focal entity serving to identify the unknown referent described by the definite noun phrase. The pseudoclefts given in (42) therefore have much in common with the predicate nominal examples given in (40).[22] Consequently, Ball's examples of reverse pseudoclefts with discourse-new information in the relative clause may not share the same function as the specificational *it*-cleft.

As I explained in §5.4, discourse-based studies of clefts, such as Prince (1978) and Collins (1991a), have found that while specificational pseudoclefts are commonly associated with expressing inferable (as opposed to discourse-old information), they do not occur with brand-new information in the relative clause. For instance, the *it*-cleft example given in (39) above, expresses a discourse-new presupposition which is not assumed to be known to the intended audience and is not inferable from the speech situation. The corresponding pseudoclefts, on the other hand, are not quite so acceptable in discourse-initial position, as shown in (43) and (44).[23] Here, the suggestion is that the audience should already know that *someone once said 'laws are silent at times of war'*.

---

[22] They also share a great deal with the Middle English examples introduced by *I it am* discussed in footnote 18. I classify this sentence type as a predicate nominal construction. Ball (1991: 41), on the other hand, labels these examples 'pseudoclefts' because of their formal similarity to reverse specificational *NP be NP* sentences.

[23] However, pseudoclefts can occur discourse-initially if the information in the relative clause is inferable from the speech situation (see Prince 1978). For example, in (i) below, the fact that the lecture is on some topic is inferable from our existing knowledge of how lectures work and is therefore present in the hearer's consciousness.

(43) (Start of lecture)
#The one who once said, 'Laws are silent at times of war' was Cicero.

(44) (Start of lecture)
#Cicero was the one who once said 'Laws are silent at times of war'.

This suggests that the development of the IP *it*-cleft requires an independent explanation and is not entirely dependent upon analogy with other specificational copular constructions.

7.2.2.4 A constructionalization story for the development of the IP *it*-cleft

I agree with Lambrecht's (1994) proposal rather than Ball's (1991, 1994a). Lambrecht argues that IP *it*-clefts do not form a separate construction from *it*-clefts with given information in the cleft clause. Instead, he assumes that these examples represent "an extension" of the *it*-cleft construction "via conventionalized pragmatic accommodation" (Lambrecht 1994: 71). Lambrecht explains that by using an expression requiring a presupposition, such as an *it*-cleft, to present information that is not shared knowledge, the expression itself forces that presupposition and the hearer accommodates it as such.

Lambrecht's (1994) hypothesis ties in well with my synchronic analysis of the present-day *it*-cleft construction. Since I assume that the relative clause and the constituent *it* function together like a discontinuous definite description, for me the *it*-cleft construction, like other definite noun phrases, must be inherently presuppositional (see §3.3.2). However, as I explained in §4.1.2.5, definite noun phrases are also associated with expressing given information. The reason for this is that in order to successfully presuppose the existence of some entity, the speaker typically has to assume that the hearer is familiar with this description. I suggest that in *it*-clefts, this general pattern of correspondence between presupposed and familiar information is overridden, allowing nonfamiliar information to be expressed in the cleft clause.

(i) (Start of lecture)
What I'm going to be talking about today is how focus can be marked in English

Historical data from the PPCME2 and the PPCEME seems to support this diachronic explanation. These corpora provide several examples of *it*-clefts in which the information in the relative clause is not discourse-old. In particular, I found 2 examples in the Late Middle English corpus where the information in the relative clause is not given by the previous discourse, but is nonetheless shared knowledge, such as (45). In this example, the proposition that *someone leant against Christ's chest* has not been mentioned before but is in the context of a discourse about the part of the Bible that this event is from. The purpose of this extract is to remind us of this scene (and so assumes that we know it), even though this particular event is not given by the previous discourse.

(45) Abid a while, I prey þe, and taak good kep ho it is þat leneþ hym so
Abide.IMP a while, I pray thee, and take good keep who it is that leans him so
boldely to Cristes brest and slepþ so sauerly in his lappe.
boldly to Christ.GEN breast and sleeps so surely in his lap.
(Aelred of Rievaulx's De Institutione Inclusarum, Vernon ms. c1400)
'Stay a while, I pray you, and take good note of who it is that leans so boldly against Christ's chest and sleeps so confidently in his lap.'

In addition, 3 examples of *it*-clefts from the LME corpus can properly be called informative-presupposition *it*-clefts. For instance, the relative clause in (46) contains discourse-new information which is not necessarily shared knowledge. However, such examples are still presuppositional, giving the impression that they express facts known to a third party. Consequently, it is perhaps unsurprising that they all belong to the genre 'history'.

(46) It was he þat graunted Kyng Herri þe Secunde to go into Yrlond and turne hem
It was he that granted King Henry the Second to go into Ireland and turn them
to þe feith,
to the faith (Capgrave's Chronicle, a1464)

By the end of the Early Modern period, we begin to see examples where the constraints are even freer. In (47), the information in the relative clause is only partly new. Here, it is given that Dunne is baulked and confused, and therefore that something *baulks thy Understanding*, but the rest of the proposition, that something *baulks thy Honesty*, is new to the discourse. In this example, the Lord Chief Justice manipulates the sense of presupposition in this construction to include his own personal opinion that Dunne is deliberately being dishonest. In this way, the use of the cleft structure allows the speaker's opinion to be presented as uncontroversial fact.

(47) Dunne. My Lord, I am so baulked, I do not know what I say myself; tell me what you would have me say, for I am cluttered out of my Senses.

L.C.J. Why, prithee Man, there's no body baulks thee but thy own self; thou art asked Questions that are as plain as any thing in World can be: it is only thy own depraved naughty Heart that baulks both thy Honesty and Understanding, if thou hast any; (Lisle, 1685)

Although the data is quite small, these examples would seem to indicate a gradual progression in the *it*-cleft construction from expressing only given information in the relative clause, to expressing shared but non-salient information, to the inclusion of information that is factual or is assumed to be factual in some sense, before finally permitting creative uses where even opinion can be accommodated into the construction. This suggests that the *it*-cleft construction has become increasingly schematic and productive over time, sanctioning new types of instances and developing a wider range of discourse functions. The new IP *it*-cleft subtype is a mismatch construction, overriding the general pattern of correspondence between presupposed and familiar information. The upshot is that particular requirements regarding the discourse status, saliency, or familiarity of the information that is accepted into the relative clause are now almost non-existent. As Borkin (1984: 125) notes for present-day *it*-clefts, "'Presupposed', when used with respect to cleft sentences, then, means 'non-asserted' or 'assumed to be true', and no more than that".

7.2.3 Summary and interim conclusions

In §7.2.2, I used data from the PPCME2 and the PPCEME to show that the non-NP *it*-cleft and the IP *it*-cleft have developed by extension from pre-existing *it*-cleft constructs. This diachronic development conforms to the principles of grammatical constructionalization outlined in §7.2.1. Over time, the *it*-cleft has undergone gradual "host-class expansion", allowing an increasing range of non-nominal foci into what was originally a strictly nominal position and accommodating new information into the presuppositional definite-like description. As a result, the *it*-cleft has developed into a more schematic and productive construction.

The historical data therefore explains how the more idiosyncratic *it*-cleft subtypes, which override inheritance from more basic constructions, have emerged gradually via conventional pathways of change. Therefore, while the *it*-cleft's construction-specific focus category and the ability for new information to be expressed in the relative clause are not inherited from more general patterns of correspondence, they are nevertheless motivated by general principles of language change. I discuss this issue in more detail in chapter 8, where I ask what the development of the *it*-cleft can tell us about the nature of constructional change and its relationship with grammaticalization.

8. THE IMPLICATIONS FOR THE PRESENT DAY ENGLISH *IT*-CLEFT

In chapter 7, I made use of historical evidence in order to provide motivation for the *it*-cleft's construction-specific properties. First of all, I showed that some of these attributes were at one time inherited from more basic constructions in the language system of earlier periods of English. Secondly, I showed that the remaining idiosyncratic properties have emerged over time as a consequence of the general principles governing constructional change. In this chapter, I examine how this historical evidence can help us to further understand the nature of the present-day *it*-cleft construction. I show how once we integrate historical evidence into the extraposition-from-NP analysis argued for in previous chapters we obtain a maximally explanatory account of the *it*-cleft construction.

In §8.1, I ask what the historical development of the *it*-cleft tells us about the nature of constructional change and go on to identify traditional diagnostics for grammaticalization in the *it*-cleft's formal and functional behaviour. From this discussion, I show how the present-day *it*-cleft construction is structured like a radial category, with more and less prototypical members. Over time, the *it*-cleft has developed extensions from the prototype; these emerging subtypes are mismatch constructions, which override general patterns of correspondence and are therefore "less prototypical". Ultimately, these new subtypes are conventionalized, which in turn has implications for the overarching *it*-cleft schema.

In §8.2, I compare the development of the *it*-cleft to that of other specificational copular constructions, drawing especially from the literature on *wh*-clefts. I argue that these related sentence-types must have undergone similar, yet construction-specific, changes which have enabled them to acquire their own unique and useful assortment of discourse functions. From this perspective, the wider specificational copular construction is also interpreted as a radial category, with each of its members radiating outwards along their own pathways of change.

Section §8.3 concludes our focus on the historical development of the *it*-cleft. Here I show how the historical evidence can be successfully integrated into the

synchronic analysis developed so far to form a maximally explanatory account of the *it*-cleft construction. I revisit, and complete, the constructional inheritance hierarchy outlined in §6.4, mapping the no longer productive inheritance relations as well as the ensuing changes which followed the emergence of less prototypical *it*-cleft subtypes. I conclude that the analysis of *it*-clefts developed throughout this thesis conforms to the principle objective of construction grammar, providing a full and explanatory account of this specialized linguistic pattern, which both tolerates construction-specific properties and maximizes motivation for the *it*-cleft construction.

8.1 The grammaticalization of the *it*-cleft construction

In this section, I ask what the diachronic development of the *it*-cleft tells us about constructional change and its integration with grammaticalization theory. I show that, in addition to the rate and direction of the change, the resulting *it*-cleft construction also exhibits well-known characteristics of grammaticalization, such as 'pragmatic strengthening' (Traugott 1988) and the 'layering' of polysemous items, whereby the original and the emergent forms coexist (Hopper and Traugott 2003: 124-126). Since most versions of construction grammar make use of default inheritance (and so assume that categories have a non-classical structure), they are well-suited to modeling the development of constructional polysemy. I conclude that a construction grammar framework is therefore a useful way of depicting grammaticalization changes to both complex and atomic (lexical) constructions.

In what follows, I reexamine the development of *it*-clefts with non-NP foci and instances with new information in the cleft clause as changes affecting the *it*-cleft as a constructional category. I show that the *it*-cleft construction is a category with a non-classical structure, in which certain instances are "better" (or more motivated) members than others. Furthermore, the *it*-cleft is structured like a 'radial category'; that is, non-prototypical members are categorized by extension from the prototype (see Lakoff 1987).[1] I suggest that the changes to the *it*-cleft construction outlined in §7.2.2 have

[1] Other constructions which have been argued to have a radial category structure include *there*-constructions (Lakoff 1987), resultatives (Goldberg and Jackendoff 2004) and ditransitive sentences

resulted in the development of these prototype-extension relations and the creation of a radial category structure for the *it*-cleft. By changing the membership of the *it*-cleft category, this in turn has consequences for how speakers conceptualize the *it*-cleft construction.

The prototypical *it*-cleft contains nominal foci and expresses given or familiar information in the cleft clause. Such examples make up the earliest attested instances of the construction and represent the most frequent subtype of *it*-cleft in present-day English. The reason for this is that this type of *it*-cleft conforms to more general patterns of correspondence in the language system. For instance, as a specificational construction involving a nominal predication relation, the *it*-cleft must contain a referring expression in the postcopular position (see §3.3 and §4.1). According to Croft (1991: 67), the semantic class of 'objects' is the typological prototype of referring constructions. Since noun phrases typically denote objects, they are therefore the phrasal category most suited to performing a referring function. Likewise, as with other definite noun phrases, the discontinuous description in *it*-clefts exhibits an existential presupposition. In order to successfully presuppose the existence of some entity, the speaker usually has to assume that the hearer is familiar with the description given. As a result, information which is presupposed is typically also given or known to the hearer.

The development of non-NP *it*-clefts and IP *it*-clefts therefore involves extension from the prototype. As I explained in §7.2.2, these emergent *it*-cleft subtypes are formed via coercion, in which non-nominal elements and information which is hearer-new are accommodated into the construction. This process is gradual since, as Goldberg (1995: 159) comments, coercion is governed by the extent to which there is relationship between the inherent meaning of the coerced item and the interpretation which it is

---

(Goldberg 1995). For example, Goldberg (1995: 31) argues that the central or basic sense of the ditransitive form involves the successful transfer of an object to a recipient, as in (i) below. However, there are also less central but related senses which do not strictly imply a successful transfer. For example, the sentence in (ii) signals only that the agent **intends** for the recipient to receive the object. Goldberg (1995: 33) concludes that the ditransitive form is therefore "associated with a set of systematically related senses"; that is, it is a radial category with a prototypical sense and non-central extensions from the prototype.

(i) Chris gave Jan a cake
(ii) Chris baked Jan a cake

given by the construction. For example, in §7.2.2.2, we saw that prepositional phrases, which often relate to nominal concepts such as time and place, are the earliest non-nominal category to occur in the focal position. In contrast, adjective phrases, which denote properties rather than objects, are not easily accommodated into the referential slot. Likewise, as I explained in §7.2.2.4, changes to the information status of the cleft clause proceed in incremental steps from expressing given information, to non-salient but shared information, to information that is factual or known to a third party, before finally accommodating the speaker's opinion. Over time then, the speaker is able to further manipulate what sort of information can be marked as presupposed or "assumed to be true".

In accordance with Himmelmann's (2004) definition of grammaticalization, the diachronic development of the *it*-cleft construction is therefore a process of gradual expansion. It involves both "host-class expansion", whereby the construction allows a wider range of components to enter into it, and "semantic-pragmatic context expansion", whereby the construction develops new pragmatic functions. Such "pragmatic enrichment" is also a diagnostic of the traditional element-based view of grammaticalization. As Traugott (1982, 1989) observes, grammaticalization often involves a shift towards increasingly subjective meanings; that is, the grammaticalized word or string of words comes to express the speaker's beliefs and attitudes. Tellingly, changes to the *it*-cleft's information structure also involve a move towards subjectification. Originally, the *it*-cleft's only function was specificational, identifying the referent that matches a familiar description. However, over time, this construction has developed additional uses, reminding the hearer of relevant but non-salient information, informing the audience of facts they may not already know and finally as an indirect way of communicating the speaker's opinion (see §7.2.2.4). The historical development of the *it*-cleft therefore involves a change in perspective from what the hearer knows to how the speaker feels.

Changes affecting the range of *it*-cleft foci and the information status of the cleft clause therefore proceed in a direction outwards from the prototype, with instances deviating increasingly from the prototype over time. Each stage in the development

therefore involves the emergence of new constructs which are "less good" and "less motivated" members of the category of *it*-clefts. These instances override general patterns of correspondence as well as overriding inheritance from the existing *it*-cleft schema. This is why default inheritance is so important for a constructional model of language change, since without it, it would not be possible for new instances to override inheritance from the *it*-cleft schema. As Hudson (2003: 366) explains, default inheritance assumes a "best-fit", rather than absolute, model of categorization and therefore allows for conflict between the inheriting and the dominating constructions.

Constructional change therefore originates in language use at the level of the construct. However, the discrepancy between the *it*-cleft schema and its less-prototypical members is resolved as the overarching category changes to suit its new membership. In other words, as the new instances become conventionalized, the speaker inductively generalizes over both the original and the emergent subtypes to form a more abstract and schematic *it*-cleft construction. As Traugott (2007: 549) comments, "If speakers adopt an innovating mismatch, by conventionalizing it, they are likely to creatively reanalyze it as a partial match that adds to the repertoire of the language". For instance, as *it*-clefts occur with non-nominal foci, the overarching *it*-cleft schema looses a syntactic constraint, with the postcopular position losing its NP specification and becoming an open slot which any phrasal category (XP) can fill. Likewise, as the *it*-cleft accommodates new information into the cleft clause, it acquires additional discourse functions which, over time, have become conventionalized uses for the *it*-cleft construction. I illustrate this stage in the development of the IP *it*-cleft by comparing examples in Early Modern English to those found in present-day English.

In the Early Modern example given here as (1), the speaker creatively exploits the presuppositional *it*-cleft construction for stylistic effect in order to present his own opinion (that Dunne is being dishonest) as uncontroversial fact (see §7.2.2.4 for a more comprehensive analysis of this example). The rhetorical device employed here is therefore "assertion by presupposition".

(1) Dunne. My Lord, I am so baulked, I do not know what I say myself; tell me what you would have me say, for I am cluttered out of my Senses.

L.C.J. Why, prithee Man, there's no body baulks thee but thy own self; thou art asked Questions that are as plain as any thing in World can be: it is only thy own depraved naughty Heart that baulks both thy Honesty and Understanding, if thou hast any; (Lisle, 1685)

While this example strikes us as an exceptional use, the stating of opinion under the guise of presupposition appears to have become a more conventionalized function for the *it*-cleft construction, as shown by the following examples from the British component of the International Corpus of English (ICE-GB).[2] Example (2) functions in many ways like a factual IP *it*-cleft, since we are told that *it was Reyner Bannon* who made a particular comment in the press. However, in this case, the factual information is presented as a follow-up to the *it*-cleft, with the relative clause expressing the speaker's approval of it.

(2) And it was Reyner Bannon who got it absolutely spot on; he commented in the press…that the architect has been driving architectural journalists mad by steadfastly refusing to release any pictures of what the Fun Palace will actually look like (S2A-040 094, 095)

Likewise, in the adverb phrase focus *it*-cleft in (3), the relative clause presents opinion as established fact. This example is particularly interesting, since *I fear* does not qualify the assertion that *others are led astray* ***here***, but the presupposition that *she risks leading others astray*.

[2] For a more comprehensive and quantitative study of present-day *it*-clefts in the ICE-GB see Gómez-González (2004), Hasselgård (2004) and Nelson (1997).

(3) And it is here, I fear, that my right honorable friend increasingly risks leading herself and others astray in matters of substance as well as of style (S2B-050 036)

Over time then, the *it*-cleft construction has gained a new function as a fairly standard, yet indirect way, of communicating, or asserting, the proposition expressed in the cleft clause (Lambrecht 1994:71). At this stage, the *it*-cleft construction is polysemous, with different types of instance conveying subtle yet distinct nuances of meaning. The development of the *it*-cleft therefore conforms to Croft's (2001: 127) constructional account of grammaticalization. He notes that, as a construction grammaticalizes, it is "extended to a new function"; over time, this new semantic use becomes conventionalized as "one of its normal functions" with the result that "the construction is polysemous with respect to its original meaning" (Croft 2001: 127).[3]

The coexistence, or *layering*, of original and emergent functions is a common outcome of the grammaticalization of lexical items, at least in its early stages (Hopper and Traugott 2003: 124-126). This is to be expected under a constructional model of language structure, which does not assume a strict division between the syntax and the lexicon. As Goldberg (1995: 31) comments, "since constructions are treated as the same basic data type as morphemes, that they should have polysemous senses like morphemes is expected". This provides support for the claim that, like lexical items, larger constructions can also be subject to grammaticalization and suggests that the construction grammar framework, which is well-suited to modeling the development of constructional polysemy, can also be usefully employed in representing grammaticalization changes to atomic, lexical items.

---

[3] Likewise, Lehmann (2008: §3.1.2) concludes that "To the extent that pragmatic accommodation of the proposition presupposed by the extrafocal clause is conventionalized in the cleft sentence, the construction becomes more grammaticalized". However, for Lehmann, the outcome of grammaticalization in this case is the "levelling out" of the contrast between the presupposition (in the cleft clause) and the assertion (in the focal position). As I have shown, presuppositional meaning is a property inherent to the definite-like description in *it*-clefts; while it may be manipulated, with the result that new information is marked as presupposed, it is not in any way reduced or lessened. As a result, I do not subscribe to Lehmann's (2008) view that changes to the *it*-cleft construction result in the "levelling out" of contrasts in information structure.

Once conventionalized, the final stage in the grammaticalization process is for the original and emergent forms to become independent from one another. For example, with atomic elements, the lexical item in its new function gains a different distribution and/or a reduced phonological form which is not shared by the lexical item in its original function. In other words, the original and emergent forms become so distinct that the speaker will not form a single abstraction over them. Croft (2001: 127) suggests that this same stage in the grammaticalization process also holds for larger, more complex constructions, whereby the construction in its new function gains formal independence. He notes that "syntactic, morphological and phonological changes…occur only to the construction in its new function, thereby making it distinct from the old construction in its original meaning" (Croft 2001: 127). However, evidence of this last step in the grammaticalization process is not found in the development of the *it*-cleft construction. For this constructional category, more and less prototypical *it*-cleft instances have the same basic structure and share the same specificational function. In other words, they are similar enough for the speaker to form a coherent schema, or higher-order abstraction.

The diachronic development of the *it*-cleft construction therefore provides good evidence for the claim that complex constructions can function not only as the surrounding context that causes the grammaticalization of an atomic component, but also as an actual grammaticalizing element (see the discussion in §7.2.1). As with other cases of grammaticalization, the *it*-cleft undergoes a series of incremental and unidirectional micro-changes, gaining a more subjective function which coexists with its original meaning. From the perspective of the constructional taxonomy, these changes originate in language use at the level of the construct. New types of instance are formed by extension from the prototype, creating a radial category structure. Both the prototype and the extension are subsumed by an overarching *it*-cleft schema, which becomes more abstract and productive as it sanctions a wider variety of different types of construct. Assuming a construction grammar model of language structure then, grammaticalization is a process of extension (or expansion) which results in (re)categorization.

8.2 A comparison with other specificational copular constructions

In §8.1, I showed that the diachronic development of the *it*-cleft involves extension from the prototype. Over time, non-NP items are accommodated into the referential slot and nonfamiliar information gains a coerced interpretation, being marked as presupposed. However, as I explained in §4.1, most types of specificational copular construction contain a referring expression and a definite NP predicate (with an existential presupposition). From this, we might expect that the different kinds of specificational sentence would undergo this same diachronic development. Nevertheless, as I explained in §5.3 and §5.4, the exact range of *it*-cleft foci is construction-specific and other specificational copular constructions cannot occur with brand-new information in the relative clause. In this section, I identify some of the subtle semantic and pragmatic properties which have instigated and shaped the *it*-cleft's construction-specific historical development. I frame my discussion in a comparison of *it*-clefts and *wh*-clefts (a topic which has received considerable interest in the cleft literature).

Despite their close familial relationship, *it*-clefts and *wh*-clefts exhibit a number of different structural and discourse-functional properties. For example, certain prepositional phrases can occur as the focus of an *it*-cleft, but are not permitted in the *wh*-cleft configuration. On the other hand, while *wh*-clefts commonly occur with verbal and adjectival foci and permit a full range of clausal foci, *it*-clefts cannot occur with verb phrase and non-factive clause foci and can only accommodate adjective phrases into the focal position under specific circumstances. Unlike *wh*-clefts, *it*-clefts can occur with brand-new information in the relative clause. Furthermore, while the *it*-cleft has developed a specialized performative function, the *wh*-cleft has acquired a construction-specific use as a presentational device.

In what follows, I explain that every one of these structural and discourse-functional differences results ultimately from two, very subtle differences in the constructions' semantic and pragmatic properties. First, we have the difference in meaning between the initial *it* of *it*-clefts and the *what* of *wh*-clefts (or *what*-clefts). As I explained in §7.1.3, the cleft *it* is semantically underspecified. This means that the head noun of the definite-like description in *it*-clefts can be given any possible interpretation.

On the other hand, while *what* has a very general meaning, it is nevertheless specified as synonymous with *the thing*.[4] The second, pragmatic difference is that, while both *it*-clefts and *wh*-clefts are associated with expressing given information in the relative clause, the two constructions nevertheless have a subtly different information structure; that is, while the information in the relative clause of *it*-clefts is typically discourse-old, that found in *wh*-clefts is normally inferable (see Prince 1978). Together, these two (related) domains of variation mean that the *it*-cleft and the *wh*-cleft have subtly different prototypes, which govern and shape their construction-specific pathways of change.

I deal with *it*-clefts and *wh*-clefts in §8.2.1 and §8.2.2, respectively. In §8.2.1, I add an additional level of detail to the diachronic account given in §8.1. Here, I focus on what caused the *it*-cleft to acquire this particular range of foci and to develop an informative-presupposition subtype. I compare this to the development of the *wh*-cleft outlined in §8.2.2. Here, I make use of the findings from both diachronic and discourse-based synchronic studies of specificational *wh*-clefts.

8.2.1 The it*-cleft's construction-specific development*

As I noted above, the initial *it* in *it*-clefts is semantically underspecified; it can be given a whole host of possible interpretations, some of which cannot be expressed by other NPs. This explains why the *it*-cleft is such a useful and productive construction and why it can sometimes contain foci which cannot easily be classified by predicative nouns. For example, the *it*-cleft in (4) is much more acceptable than the corresponding *wh*-cleft and *th*-cleft given in (5) and (6).

---

[4] The distinction is made particularly apparent in predicational clefts. As Ball (1991: 58) notes, predicational *it*-clefts often have a "more restrictive interpretation" than predicational *what*-clefts. For example, while the *it*-cleft in (i) in understood to mean *the dress that she wore was wonderful*, the *what*-cleft in (ii) tells us that *the thing she wore was a wonderful dress*. In this example then, the initial *it* in *it*-clefts is interpreted as having a much more specific meaning than *the thing*, gaining its semantic interpretation from the postcopular noun (see §5.2.2). This shows that the cleft pronoun *it* is therefore underspecified, rather than truly general, preferring a specific (or restrictive) interpretation.

(i) It was a wonderful dress that she wore
(ii) What she wore was a wonderful dress

(4) It was with great determination that he climbed to the top

(5) *How he climbed to the top was with great determination

(6) *The way that he climbed to the top was with great determination

Here, the prepositional phrase *with great determination* does not quite fall within the semantic scope of *how* or the adverbial noun *way*, which typically denote 'instrument' or 'means'. As a result, in these examples, the referring expression does not match up with the head noun of the definite description. On the other hand, since the cleft *it* is semantically underspecified, it covers "the same ground as the adverbial nouns **and much more**" (Bolinger 1972: 122; emphasis added).[5]

However, while the cleft *it* is maximally underspecified, the interpretation it is given is usually specific. In other words, the discontinuous constituent in *it*-clefts uniquely describes the postcopular referent. This explains why 'listing' is an effective coercion strategy for accommodating items which denote properties into the referential slot (see §7.2.2.2). As Declerck (1984b: 144) comments, "properties are not mutually exclusive: if X has property A, there is no reason why it should not have other properties as well". For example, in (7a) and (8a) *the thing that he was* and *the thing that he is* do not uniquely describe the adjective phrase *sick* and the predicative noun *a secretary*. However, by contrasting *sick* with *tired* in (7b), the property *sick* is established as a member of a restricted set of possible ailments. Likewise, in (8b), the predicative NP *a secretary* is contrasted with the profession that uniquely matches the more restrictive description *the thing that I'd wanted to be*.

(7) a) *It's sick that he was

b) It's not sick that he was but tired (É. Kiss 1998: 262)

---

[5] A similar explanation can be provided to account for the difference in acceptability between the *it*-cleft in (i) and the corresponding *th*-cleft in (ii). While for Huddleston (1984: 461) these examples are problematic for an extraposition-from-NP analysis of *it*-clefts, they are easily accommodated into our account. Although the prepositional phrase *to Ed* cannot be classified by the head noun *person*, it can nevertheless be described by the semantically underspecified cleft pronoun.

(i) It was to Ed that she was referring (Huddleston 1984: 460)

(ii) *The person that she was referring was to Ed

(8) a) *It's a secretary that I am

b) They made me a secretary, but it wasn't a secretary I'd wanted to be

(Ward, Birner and Huddleston 2002: 1418).

Nevertheless, not all predicative elements can be 'individualized' in this way. For example, restricting the set description and listing possible alternatives in (9b), does not make the verb phrase *eat* much more acceptable than in (9a). For the semantic class of 'actions' to fill the referential slot of the *it*-cleft, they must become more noun-like, as with the gerund *eating* in (9c). This example is closer to the prototypical *it*-cleft instances, which contain nominal foci denoting discrete entities (see §8.1).

(9) a) *It's eat that she does

b) *It's eat that she likes to do best, not shop

c) It's eating that she likes best, not shopping.

In addition to the restrictions on verb phrase foci, *it*-clefts cannot occur with non-factive clauses in the postcopular position (see Delahunty 1984; Delin 1989). For example, (10a) contains the non-factive verb *say*. Here, the proposition *that he never replied to her letters* is not assumed to be true since it is attributed to the thoughts and beliefs of the person who said it and could therefore involve conjecture or falsehood. (10b), on the other hand, contains the factive verb *regret*. Here, the clausal complement is understood to express a true proposition. In other words, *he regretted **the fact** that he never replied to her letters*.

(10) a) *It was that he had never replied to her letters that he said

b) It was that he had never replied to her letters that he (most) regretted

Again, the reason for this difference in acceptability seems to lie in the fact that sentences such as (10b) are closer to the *it*-cleft prototype. Here, the factive clause refers to an actually occurring event. In such sentences, the *that*-clause can be replaced by

other noun-like elements, such as gerunds (*his insistence that he was in the right*) and adjectival nominalizations (*his carelessness*) (see Kiparsky and Kiparsky 1970). However, these nominal constructions cannot replace the *that*-clause complement of non-factive verbs. It makes sense then, that when presented with *that*-clause foci, the hearer automatically interprets them as referring to existentially presupposed events, rather than to some abstract concept such as *the words somebody said.* Therefore, in example (10a) above, the non-factive verb *say* conflicts with our prior assumption that the proposition expressed in the postcopular clause is true.

As I explained in §8.1, the *it*-cleft prototype is associated not only with NP foci, but also with expressing given information in the cleft clause. What is more, the information in the relative clause is normally discourse-old. As a result then, the *it*-cleft construction is understood as providing the unknown referent that uniquely matches a pre-existing description. In other words, as Hedberg (1990: 123) observes, *it*-clefts have the function of providing the correct "answer" to an already established "question". In support of this, Bolinger (1977: 71) finds that, in *it*-clefts, the specifying relationship between the focal element (or 'value') and the definite-like description (or 'variable') must have a prior basis. He provides the following examples which differ in acceptability.

(11) A: When will we know?
B: It's tomorrow that we'll know
(12) A: When will you tell me?
B: #It's tomorrow that I'll tell you (Bolinger 1977: 71)

The discourse in (11) would be entirely appropriate if speakers A and B are waiting for exam results which are to be made available on tomorrow's date. As a result, in this example, the time of knowing has been previously established. In (12), on the other hand, the time of telling has no prior basis and is decided by the speaker there and then. Consequently, an *it*-cleft is not the best choice for expressing this information.

The early association between *it*-clefts and the presentation of pre-established facts provides a plausible explanation as to why the *it*-cleft, as opposed to any other kind of specificational sentence, has developed an informative-presupposition subtype. As I explained in §7.2.2.4, the IP *it*-cleft expresses new, as opposed to given, information in the cleft clause. Nevertheless, such examples have the function of expressing this information as uncontroversial fact, signaling for the hearer to accept it unchallenged (see Delin 1992). The *it*-cleft configuration therefore provides a useful way to reduce the speaker's responsibility in the information expressed; as Prince (1978: 900) comments, IP *it*-clefts "function like traditional footnotes in that they seem to say: 'Don't argue with me – I didn't invent this – and I'm aware that I didn't invent this'".

This idea, that *it*-clefts have the effect of downplaying the speaker's role, also helps to explain why the *it*-cleft has developed a performative function, which is unique among the family of specificational copular constructions. According to Prince (1978: 903), examples such as (13) below downplay the speaker's position of power, while at the same highlighting their emotional response (*with great pleasure*).[6]

(13) It is with great pleasure that I present to you this award

[6] Los (2009) interprets the function of performative clefts differently. She uses these examples to build up a case that the *it*-cleft has developed in response to the loss of verb-second in English. Los (2009: 111) claims that the resulting SV word order has "compromised the ways that were available to the language user to structure information". One such effect is that the pre-subject position (involved in preposing) has become pragmatically marked as expressing prominent information. According to Los, the *it*-cleft provides us with a useful strategy to avoid positioning adverbials in too prominent a position. She provides the following examples which differ in acceptability.

(i) #With great pleasure, we can inform you that your application was successful
(ii) It is with great pleasure that we can inform you that your application was successful
(examples from Los 2009: 114)

Since I view the *it*-cleft primarily as a specificational copular construction, I am skeptical of Los' wholly information-structural diachronic story. Furthermore, Los' analysis is fundamentally flawed. She claims that the function of the IP *it*-cleft in (ii) is to "place *with great pleasure* in end-focus position to make it *less* marked" (Los 2009: 114; italic original). Of course, this is paradoxical, since the *it*-cleft is a focusing construction and is not a neutral way of presenting information. Although Los (2009: 114) recognizes this, she claims that "nothing else appears to explain the awkwardness of the more literal translation".

In this instance, the actual *presentation of the award* is backgrounded, or taken for granted, in order to show deference or politeness to the addressee. Since this construction is so commonplace in ceremonial contexts, the (now) formulaic device also acts as an anticipatory signal to the audience that the ritual has begun.

8.2.2 The what*-cleft's construction-specific development*

As I explained in §7.2.2, *it*-clefts were originally restricted to occurring with noun phrases in the focal position. This supports the claim that examples with NP foci make up the prototypical *it*-cleft subtype. However, according to Traugott's (2008) recent exploratory work, the early *wh*-cleft instances do not share this same property. She finds that the specificational *wh*-cleft did not emerge until the late 17th century, at a time when other specificational copular constructions, including the *it*-cleft, were already well established. From her data, it appears that the early *what*-cleft could occur not only with NP foci, but also with a range of clausal foci (including both factive and non-factive clauses) and verb phrase foci (such as *to*-infinitives).[7]

From the very beginning then, the *wh*-cleft could occur with categories of foci which are not acceptable in the *it*-cleft. For example, the *wh*-cleft in (14), containing a non-factive clause focal element, is perfectly grammatical.

(14) What he said was that he had never replied to her letters

There are two possible reasons for this. First, since *wh*-clefts are not associated with nominal foci, the *that*-clause will not necessarily be interpreted as a noun-like concept, such as an actually occurring event (see §8.2.1). Second, *wh*-clefts have a different linear

[7] Traugott (2008) suggests that *wh*-clefts with infinitival foci may have developed from non-specificational tokens. Originally, examples such as (i) had only a purposive meaning (*the thing I do is in order to please you*). With the emergence of the specificational *wh*-cleft, such examples could be reanalysed. On this new reading, the *to*-infinitive is identified as the *thing that I do* rather than *my purpose in doing it*. The independence of this new meaning from the original purposive construction is accomplished by the fact that *to* finally becomes an optional element, allowing bare infinitive foci, shown in (ii).

(i) What I do is to please you
(ii) What I do is please you

order to *it*-clefts, which contain extraposed relative clauses. In example (14) then, the non-factive verb *said* is given prior to its clausal complement, marking it as non-presupposed. As a result, the hearer cannot misinterpret the focal clause as factive.[8]

The reason why *wh*-clefts are more open to non-nominal categories than *it*-clefts lies in the semantic difference between the constructions' introductory elements. While the initial *it* in *it*-clefts is semantically underspecified, the *what* of *what*-clefts is specified as synonymous with *the thing*. In the *wh*-cleft, this very general concept is extended to cover not just objects, but also actions and properties. As I explained in §8.2.1, these sorts of entities are not mutually exclusive and as a result, the noun phrase headed by the semantically general *what* does not always provide a description which is unique to the postcopular referent. For example, in (15), (16) and (17) below, there is not only one possible thing that *John is* or *Sarah was* or that *she can do*. Nevertheless, these property-denoting foci can be accommodated into the *wh*-cleft without requiring the use of coercion strategies, such as listing (see §8.2.1).

(15) What John is is stupid

(16) What Sarah was was an idiot

(17) What she can do is eat

In *wh*-clefts then, there is not always the sense that we are giving the correct "answer" to a specific "question". Instead, in this construction, there is a certain flexibility which allows the speaker to choose between different, but equally correct, alternatives. As a result of this, the *wh*-cleft emphasizes the speaker's role in the specificational process. For example, in (15) and (16) above, the speaker selects the characteristic which they believe appropriately sums up the individual in question. In doing this, the speaker is able to express their own opinions about that individual.

---

[8] The linear order of the predicative and referring elements cannot be the only reason for the grammaticality of *wh*-clefts with non-factive foci. Otherwise, we could not explain why reverse *wh*-clefts are deemed slightly more acceptable than their corresponding *it*-clefts (see also Delin 1989: 97).

(i) ?That he had never replied to her letters was what he said

Likewise, in (17), the speaker selects an activity which *she does well*. By highlighting one particular activity over many possible others, the speaker is able to make a comment about the individual in question, for example, that *she is greedy*.

Therefore, unlike the *it*-cleft, the *wh*-cleft is not associated with presenting factual, or pre-established, information. This explains why the *wh*-cleft has never developed an informative-presupposition subtype. In such sentences, new information is given the status of fact; that is, the *it*-cleft construction is employed for the purpose of reducing the speaker's responsibility for the information in the cleft clause. However, as we have seen, *wh*-clefts have acquired a very different function, which highlights the speaker's role in the act of specification. Furthermore, unlike the *it*-cleft, the *wh*-cleft construction is not associated with expressing discourse-old information in the relative clause. For instance, Traugott (2008) finds that, from the very beginning, *wh*-clefts were associated with inferable information, just as they are today (see §5.4). As Prince (1978) comments, information is 'inferable' if it is appropriate to the speech situation and can therefore be assumed to be either already in the hearer's consciousness or easily constructible from the discourse context. The function of *wh*-clefts then, is not to provide the correct "answer" to an already established (or discourse-old) "question". Instead, as Hedberg (1990: 123) comments, the *wh*-cleft is used "to both ask and answer a question".

Such properties explain why the *wh*-cleft has developed a presentational function. In these examples, the focal element expresses the main informational content of the sentence and the definite-like description simply functions as a presentational device. For instance, in (18), the initial description and the copular verb function as a unit which introduces the postcopular proposition.

(18) What I'm saying is, you shouldn't let her boss you about

Here, the initial description is inferable, since, according to Prince (1978: 891), "the speaker's relevant thoughts, observations [and] opinions…are taken to be the constant appropriate concern of the hearer". Because of this, the initial description, and indeed

the specifying relationship between it and the postcopular referent, is secondary to the proposition expressed in the focal position. Tellingly, the clausal focus is presented here as a complete sentence; that is, unlike the clausal foci of *it*-clefts, it does not have to be introduced by *that*.[9] Bolinger (1977: 11) notes that the word *that* is inherently anaphoric. The lack of *that* in (18) therefore indicates that the postcopular clause is a brand-new assertion. Prince (1978: 891) says, "such WH-clefts are often used not simply for clarifying previous assertions, but also for remaking them".

In these presentational *wh*-clefts, the proposition expressed in the focal clause is more informative than the specifying relationship between the focal element and its description. According to Koops and Hilpert's (2009) diachronic study, the *wh*-cleft has come to be introduced by increasingly general descriptions over time. For example, while the highly general verb *do* was often found in the precopular phrase of early *wh*-clefts (see also Traugott 2008) they are now also common with the even less specific verbs *happen* and *be*. Koops and Hilpert (2009) note that while the verb *do* requires that the agent of the dynamic event is specified (*what **he** did was...*), the verb *happen* does not (*what happened was...*); in turn, the verb *be* is "maximally general" since it can encompass both dynamic and stative events (*what it was was...*).

According to Hopper (2001), presentational *wh*-clefts play a significant role in conversational turn-taking. He notes that the precopular description serves to delay the delivery of the main assertion, "by adumbrating (foreshadowing) the continuation in general terms without giving away the main point" (Hopper 2001: 114). This has the dual purpose of informing the reader that what follows is worthy of attention as well as buying time for the speaker to formulate the postcopular assertion, therefore enabling

---

[9] Such examples provide support for the claim made by Den Dikken, Meinunger and Wilder (2000) and Schlenker (2003) that NPI connectivity in *wh*-clefts can be accounted for by invoking ellipsis. The canonical *wh*-cleft is the only specificational construction that licenses negative polarity items in the focal position, shown in (i). As I explained in §4.3, if we assume that such sentences are in some way related to *wh*-clefts with sentential foci, such as (ii), then NPI connectivity in *wh*-clefts is explained. However, since examples with sentential foci are a later development, forming a separate 'presentational' subtype, an ellipsis approach is not a valid analysis for all *wh*-cleft instances, *contra* Schlenker (2003).

(i) What he didn't buy was any wine
(ii) What he didn't buy was he didn't buy any wine

them to "hold the floor".[10] In *wh*-clefts then, the specificational relationship between the precopular description and the postcopular referent is often completed "online", decided by the speaker during the course of the utterance. Hopper (2001) goes on to show that in some cases, this specifying relationship is actually left incomplete.

Over time then, the *wh*-cleft has developed a new discourse function in which the act of specification is so general that it functions only as a simple presentational device. Indeed, for some speakers it seems that the construction no longer has a specificational meaning at all. This explains the purpose of the "double *is*" construction, exemplified by (19), which contains two instances of the matrix copula.

(19) What it is **is, is** that I just can't see the point of doing it

In such sentences, the initial description (in this case, *what it is*) functions as a unit with the first matrix copula *be* (see Brenier and Michaelis 2005). Since this initial unit of information functions only as a formulaic presentational device, a further copula is provided after the intonation break to reinforce the specifying relationship between the postcopular clause and the initial description. This construction has a further turn-taking function, providing even more "down time" for the speaker to formulate their main assertion.

8.2.3 Summary and interim conclusions

Throughout §8.2, I have shown that while the *it*-cleft and the *wh*-cleft are both members of the family of specificational copular sentences, and therefore require a similar analysis (see §4.1), they have nevertheless undergone construction-specific diachronic developments, expanding in different directions. From the very beginning, the *it*-cleft construction was associated with presenting factual information, whereby the correct referent is identified as matching a specific, and discourse-old, description. This explains why the *it*-cleft construction developed an informative-presupposition subtype, in which

[10] See also Schmid (2001: 1536) on the use of definite NPs with abstract head nouns, such as *the thing is*, as "a useful hesitation device".

new and even subjective information is given the status of fact. In contrast, the *what*-cleft extends to encompass all those semantic categories that can be classified as *things*, including actions and properties. Since such elements do not represent discrete, mutually exclusive entities, the descriptive component of the *what*-cleft is highly general, leading to a new discourse function as a presentational device. Therefore, despite the fact that *it*-clefts and *wh*-clefts inherit from the same basic construction, they have nevertheless developed their own unique radial category structures[11]. This, in turn, serves to make the overarching specificational copular schema a more open (or schematic) and productive category.

8.3 An updated inheritance hierarchy

As I explained in §2.2, construction grammar assumes that a speaker's linguistic knowledge is made up of constructions which are organized within a hierarchical network. By examining how one construction relates to other similar constructions, the linguist is effectively showing how a small subsection of the speaker's grammatical network is most likely organized. In §6.4, I sketched a diagrammatic representation based on the observations made and the analyses developed throughout chapters 3, 4 and 5 (see Figure 6.3). This hierarchy of copular constructions illustrates the inheritance links through which *it*-clefts are related to other more schematic constructions. In particular, it shows that the different types of *it*-cleft inherit their specificational and predicational meanings (both of which involve the same classifying, nominal predication relation) from the wider canonical specificational and predicate nominal constructions.

However, as I noted in §6.4, this interim inheritance hierarchy is incomplete. For one thing, it does not explain how the *it*-cleft configuration, involving a restrictively modified pronoun and an extraposed relative clause, ever came into being. Since these structural properties are specific to the *it*-cleft construction, they are not supported by inheritance from the present-day English language system. As Goldberg (2003: 120-1) comments, if we cannot explain why a construction should exist in the language, then

[11] Croft and Cruse (2004: 319) point out that distinct constructions often emerge from specific instances of existing constructional schemas, expanding in their own directions. See also Israel (1996) on the *way*-construction.

we do not yet have a fully explanatory account of that construction. However, once we incorporate the historical evidence from §7.1, our inheritance hierarchy gains a diachronic dimension and the *it*-cleft's more irregular structural properties are identified as the historical remnants of once regular patterns.

As I explained in §7.1, the analysis of *it*-clefts outlined in §4.1 is supported by the historical evidence; that is, it is in keeping with what we know about the language system of earlier periods of English. For example, in this thesis, I have argued that the *it*-cleft construction contains a nominal predication relation with the initial *it* and the sentence-final relative clause functioning together as a definite NP predicate. Although restrictive relatives cannot normally modify pronouns, this linguistic pattern was a much more general phenomenon in Old and Middle English (see §7.1.1). At earlier stages in the *it*-cleft's history then, the restrictive modification of the pronoun *it* was motivated by, and inherited from, the no longer productive determinative pronoun construction.

In the account of *it*-clefts argued for in this thesis, the cleft clause is therefore analysed as an extraposed restrictive relative. As I explained in §4.3, those who criticize extraposition accounts often ask why this restrictive relative clause cannot occur in a position adjacent to its pronominal head. Again, the historical evidence provides us with the answer. While the extraposition of a restrictive relative clause is atypical in present-day English, relative clauses were often found sentence-finally in Old English (see §7.1.2). Rather than deriving from non-extraposed configurations, many authors claim that the OE relative clause originated as a paratactic structure (see Ball 1991: 60). For the early *it*-cleft then, the lack of a non-extraposed variant is to be expected. While the relative clause construction has undergone important changes, with present-day instances typically forming syntactic units with their antecedents, the *it*-cleft construction has remained unchanged (possibly due to information-structural and/or prosodic factors (see §7.1.2)).

Historical evidence also helps to explain the behaviour of the cleft pronoun. In my analysis of *it*-clefts, the initial pronoun *it* is not semantically empty. Instead, it performs an important quantifying role and functions as the head noun of a definite-like description. However, the present-day English lexeme *it* is inherently singular and it

marks the referent as non-human. How then can we explain the occurrence of *it*-clefts with human and/or plural foci? Well, as I explained in §7.1.3, the lexeme *it* in Old and Middle English is maximally underspecified for both number and animacy. Therefore, although the cleft pronoun is now morphologically singular, it nevertheless retains the original semantic properties of the Old English lexeme. This, in turn, provides support for the explanation of the *it*-cleft's unusual number agreement patterns outlined in §4.3. Here, I claimed that while the matrix verb agrees with the singular set denoted by the pronoun *it*, the verb embedded in the cleft clause shows agreement with the membership of this set (which may be plural or singular).

At the point of origin then, the *it*-cleft configuration was fully motivated by the language system of the period, inheriting properties from the determinative pronoun construction, the (paratactic) relative clause construction and the lexeme *it* (see Figure 8.1). While these constructions have either fallen out of productive use or have undergone important changes, their influence remains entrenched within the *it*-cleft's now idiosyncratic structure.[12] As a result, in Figure 8.1, these inheritance relations are depicted by dashed lines, indicating that they are representative of an earlier stage of the language.

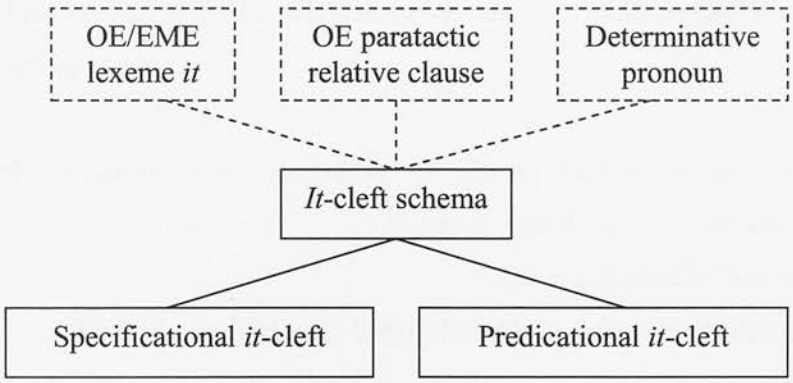


Figure 8.1 *Motivation for the* it*-cleft configuration: no longer productive inheritance relations*

[12] In order to model this diachronic development, we need a constructional inheritance hierarchy in which information is stored redundantly. In this model, the information inherited from the dominating construction is also specified in the inheriting construction (see §2.2).

The historical evidence therefore provides motivation for the *it*-cleft's construction-specific structural properties; that is, it provides an explanation for the construction's existence in the language. In essence then, this evidence also provides us with an origin story for the *it*-cleft. We might speculate that the Old English *it*-cleft developed from instances of the canonical specificational construction with the pronominal subject *it*. For example, in the specificational *NP be NP* sentence in (20), the initial *hit* is a full anaphoric NP with the meaning *the one that was standing there.*

(20) *…sæde þæt Petrus þær stode. þa geleaffullan cwædon þæt hit nære Petrus,*
…said that Peter there stood the faithful said that it not-were Peter
ac wære his engel. (*Ælfric, Catholic Homilies*, vol. II, 382.21)
but were his angel
'[Rhoda] said that Peter was standing there. The faithful said that it wasn't Peter, but was his angel.' (Ball 1991: 24)

Once we add a paratactic relative clause (in accordance with the determinative pronoun construction), we obtain a specificational *it*-cleft, as in (21). Here, the initial *hit* is restrictively modified by the sentence-final clause and functions as the definite article and the underspecified head noun of the definite-like description *it (the one) that is knocking there.*

(21) *þa cwædon þa geleafullan, 'Nis hit na Petrus þæt þær cnucað, ac is his ængel.'*
Not-is it not Peter that there knocks but is his angel
(*Ælfric, Catholic Homilies*, vol. I, 517-18.1)
'Then the faithful said: It isn't Peter who is knocking there, but his angel.'
(Ball 1991: 39)

The historical evidence therefore suggests that the *it*-cleft originated by extension from the already existing canonical specificational *NP be NP* construction. As Ball (1991: 24)

comments "it is plausible that a language could have the simple copular sentence without having the cleft".

By incorporating the historical evidence from §7.1, the inheritance hierarchy outlined in Figure 6.3 therefore gains a diachronic dimension, depicting inheritance relations which existed at earlier periods of the language (see Figure 8.1). However, in some ways Figure 6.3 already illustrates an earlier stage of English. For example, the hierarchy does not cover non-NP *it*-clefts or informative-presupposition *it*-clefts – subtypes which only began to emerge hundreds of years after the *it*-cleft's origin. In what follows, I show how once we incorporate the diachronic data from §7.2, the inheritance hierarchy gains an added complexity, representing a maximally explanatory account of the *it*-cleft construction.

In §7.2, I showed that the informative-presupposition *it*-cleft emerges by extension from the existing *it*-cleft prototype. I argued, in §8.2, that because the earliest instances of this construction are associated with expressing factual information (with the proposition in the cleft clause having a discourse-old information status), the construction gradually acquires a new discourse function. Over time, hearer-new information is incorporated into the cleft clause and is marked as uncontroversial fact. This historical development is illustrated in Figure 8.2. Here the specificational *it*-cleft at t1 (indicated by dashed lines) is extended to form a new sub-construction: the IP *it*-cleft. The speaker abstracts over the instances sanctioned by the specificational *it*-cleft at t1 and the instances which make up the new IP *it*-cleft (shown by the two upward arrows). As a result, the specificational *it*-cleft has become a more schematic and productive, and therefore a higher-order, construction.

Finally, the temporary inheritance hierarchy outlined in Figure 6.3 assumes that all of the copular constructions depicted are *NP be NP* sentences. There are several reasons why we began with this assumption. First, the cleft literature centres on examples with NP foci, which are by far the most common subtype in present-day English. Secondly, the literature on specificational copular sentences treats them as having an *NP be NP* configuration (see §3.2.2). In addition, these more prototypical examples helped to clarify the inheritance relation between *it*-clefts and the canonical

specificational construction as well as the relation between predicate nominal sentences and their "inverse" (canonical specificational) counterparts.

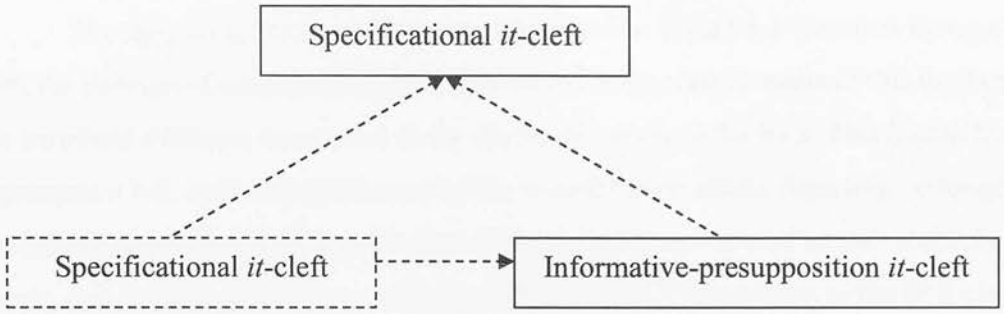


Figure 8.1 *The emergence of the IP* it-*cleft and the schematization of the specificational* it-*cleft*

In §7.2, I showed that *it*-clefts with non-NP foci originated by extension from the existing NP-focus *it*-cleft. However, unlike the IP *it*-cleft, which seems to have developed a unique discourse function, it is unlikely that non-NP *it*-clefts form a distinct construction (see Figure 8.3). Instead, I assume that the speaker abstracts over instances with NP foci as well as the instances with non-NP foci, with the specificational *it*-cleft becoming a more schematic construction; that is, the *it*-cleft changes from an *It be NP relative clause* construction to an *It be XP relative clause* construction. However, while the focal slot is syntactically underspecified, the range of possible foci is dependent upon the semantic referentiality requirement.

Once we bring non-NP *it*-cleft instances into the inheritance hierarchy, we can no longer assume that the *it*-cleft inherits from a higher-order *NP be NP* construction. Instead, and as shown in Figure 8.3, the specificational *it*-cleft inherits from the canonical specificational construction which has a *subject be complement* syntax and a 'classifying' semantic predication relation (containing an XP referring expression and a predicate nominal).[13] Therefore, the updated inheritance hierarchy in Figure 8.3 does not provide any restrictions on the range of elements that can fill the referential (XP)

[13] It is possible that speakers abstract over the canonical specificational and predicate nominal constructions, recognizing that they contain the same semantic (class-membership) predication relation. However, as Croft and Cruse (2004: 308) note, speakers "may not have the most schematic constructions represented in their minds". As I result, this level of abstraction is not represented in Figure 8.3.

position of these copular constructions. Instead, this is understood to be dependent upon construction-specific factors which are reflected in the construction's own historical development (see §8.2).

The updated inheritance hierarchy illustrated in Figure 8.3 therefore incorporates both the patterns of correspondence identified in the synchronic parts of this thesis and the historical evidence uncovered in the diachronic study of *it*-clefts. This hierarchy represents a full, explanatory account of the *it*-cleft construction, depicting no longer productive inheritance relations (indicated by dashed lines) as well as the relatively recent, and increasingly idiosyncratic, *it*-cleft subtypes.[14] In addition to the IP *it*-cleft, the performative *it*-cleft, the proverbial *it*-cleft and the presentational *wh*-cleft are represented in Figure 8.3 (and indicated in bold). Although these constructions are unproductive and formulaic, since they are made up of a limited set of highly entrenched instances, they nevertheless have distinct discourse functions and therefore need to be recognized as constructions in their own right.

In conclusion then, while the interim hierarchy in Figure 6.3 emphasized the familial correspondences between *it*-clefts and other copular constructions, the updated hierarchy in Figure 8.3 also represents the idiosyncratic information which cannot obtain a satisfactory explanation on a purely synchronic account. Over the last two chapters, I have shown that while a number of the *it*-cleft's properties are construction-specific, they were nevertheless at one time motivated either by the language system at earlier periods of the language or by the general principles governing language change. As I explained in §6.3, a constructional approach is one which tolerates idiosyncratic information but nevertheless aims for explanatory adequacy. By undertaking a synchronic and diachronic analysis of the *it*-cleft construction, I have provided a full account of the specialized linguistic pattern, which both tolerates construction-specific properties and maximizes motivation for the *it*-cleft construction.

[14] However, in some ways, the hierarchy of copular constructions outlined in Figure 8.3 is still incomplete. For simplicity, I have omitted demonstrative clefts, *there*-clefts and specificational sentences with indefinite NP predicates from this taxonomic network of constructions.

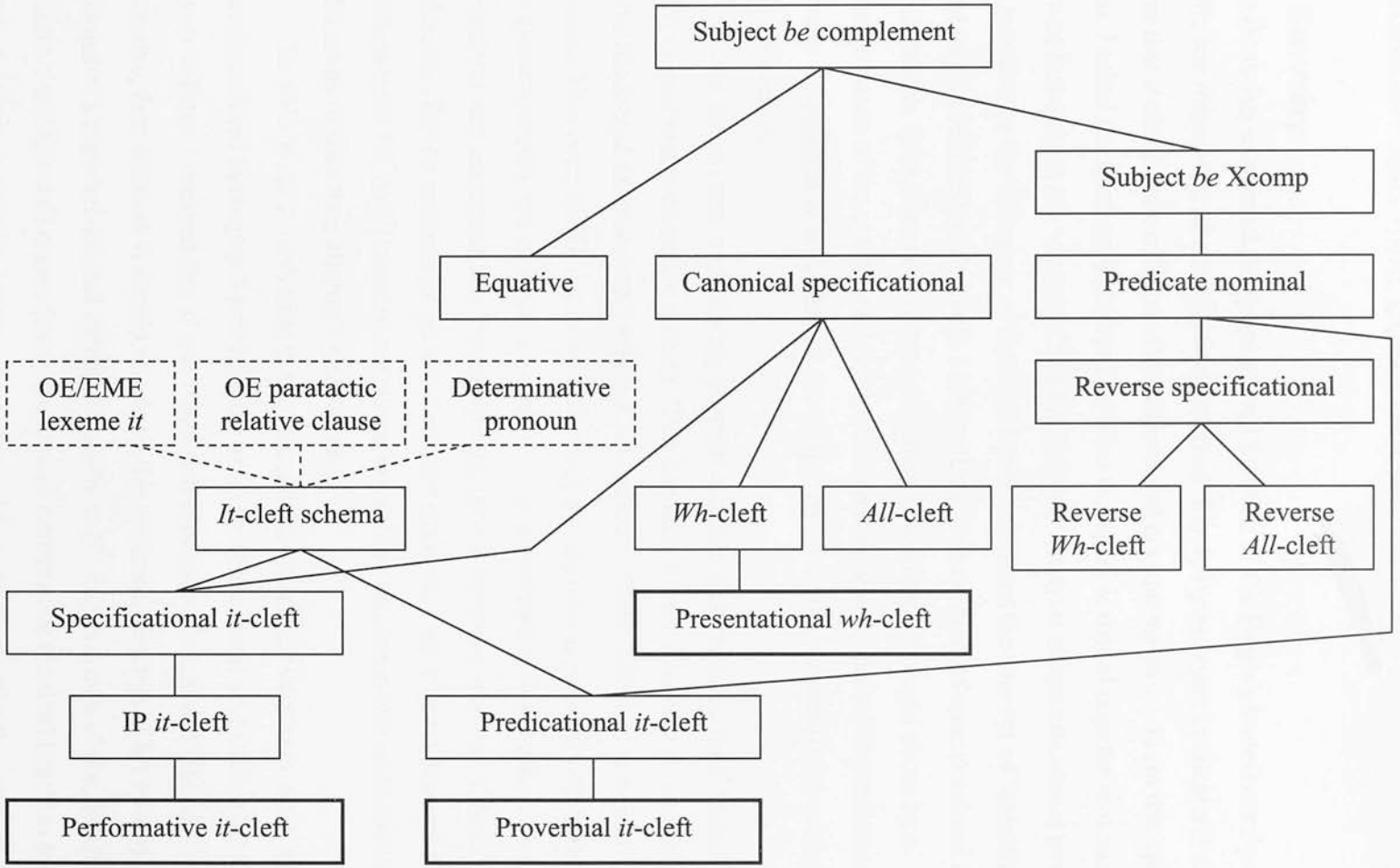


Figure 8.3 *A constructional taxonomy of the* it-*cleft's inheritance relations*

9. SUMMARY AND FINAL CONCLUSIONS

9.1 Summary

This thesis has examined the structure and function of the English *it*-cleft configuration within the framework of construction grammar. My analysis began in chapter 3 with the claim that *it*-clefts are a subtype of specificational copular sentence. From this starting point, I asked the fundamental question: what is a specificational copular sentence? I showed that neither the 'equative' nor the 'inverse' analysis of specificational structures can account for the full range of data and I concluded that the concept of 'specification' is not well understood. As a result, I outlined my own analysis of specificational *NP be NP* sentences. Here, I argued that specificational meaning is brought about by a reinterpretation of the class-membership relation involving definite NP predicates, whereby the referent is identified as the unique member of a restricted and existentially presupposed set.

Like the inverse analysis, my account assumes that specificational sentences involve a nominal predication relation. This accounts for the asymmetric interpretation of specificational sentences and explains Mikkelsen's (2002, 2005) pronominalization evidence. However, unlike the inverse account, my analysis is primarily semantic rather than syntax-centred and so is not dependent upon movement or linear order. As a result, my analysis can accommodate reverse specificational sentences and provides an explanation for the restrictions on indefinite NP subjects. This original account of specificational *NP be NP* sentences is therefore able to accommodate and explain a wider range of data than alternative analyses.

In addition to contributing to the literature on specificational sentences, the analysis outlined in chapter 3 provided the basis for the account of *it*-clefts given in chapter 4. Here, I claimed that if specificational sentences contain definite NP predicates, then the *it*-cleft should also have this semantic structure. This led me to argue in favour of a non-derivational extraposition-from-NP analysis of *it*-clefts, in which the pronoun *it* and the cleft clause (analysed here as a restrictive relative) function together as a discontinuous definite description. I showed how, once we take this step, many of

the *it*-cleft's seemingly idiosyncratic properties are inherited from the wider specification construction and can be explained either as resulting from the semantics of definite noun phrases or as a product of specificational meaning. I concluded that the analysis put forward in this thesis, in which the initial *it* performs a definite-like quantificational role, therefore has much more explanatory power than the 'expletive' accounts of *it*-clefts, in which *it* is analysed as a semantically inert element.

The particular analysis of *it*-clefts argued for in chapter 4 improves upon the existing 'extraposition' and 'discontinuous constituent' accounts that have been proposed in the cleft literature. Most importantly, my analysis is integrated into an original account of specificational sentences which recognizes the crucial role that definite descriptions play in creating specificational meaning. It is this that provides the most compelling support for a discontinuous constituent analysis of *it*-clefts. Furthermore, while several authors have identified patterns of correspondence between *it*-clefts and definite noun phrases, I have applied classic philosophical work on definite descriptions to the *it*-cleft data, thereby explaining exactly how these properties come about.

In chapter 5, I established the scope of my unified analysis of specificational copular constructions. I showed how, once we analyse specificational sentences as involving nominal predication, the relationship between specificational and predicational *NP be NP* constructions, and consequently the relationship between specificational and predicational *it*-clefts, becomes straightforward. Since predicational *it*-clefts are particularly problematic to expletive *it*-cleft analyses, this represents an important advantage to my analysis. Finally in chapter 5, I acknowledged the limits of my unified account, showing that the range of *it*-cleft foci and the information status of the cleft clause are construction-specific properties which require an independent explanation.

In chapter 6, I compared my account to other constructional *it*-cleft analyses proposed in the literature. Here, I showed that my account is more in keeping with the basic claims that underlie the constructional framework and makes better use of the machinery that provides a constructional approach with explanatory power. In addition to serving as an interim summary of my analysis, this discussion contributes to the

literature on construction grammar. It illustrates how the constructional framework is often misinterpreted as promoting highly idiosyncratic and *ad hoc* analyses (by both critics and supporters alike) and outlines the true advantages of a constructional approach for work on specialized linguistic patterns.

The analysis developed in chapters 3, 4 and 5 focused on maximizing motivation for the *it*-cleft construction via inheritance from higher-order constructions in the language system. While this account was shown to have explanatory adequacy, certain aspects of the *it*-cleft's structure and use remained unexplained. In chapter 7, I made use of historical evidence in order to find motivation for the *it*-cleft's more idiosyncratic properties. I found that (a) much of the *it*-cleft's structure is reminiscent of an earlier stage of the language and (b) the *it*-cleft's construction-specific range of foci and discourse functions have emerged over time as a result of general principles of language change. This chapter contributes to the somewhat limited literature on the history of the English *it*-cleft, providing an original account of the specificational *it*-cleft's origins and its subsequent diachronic development.

In chapter 8, I showed how the historical evidence can help us to further understand the nature of the present-day *it*-cleft construction and its relationship to other types of specificational copular sentence. As part of this discussion, I asked what the development of the *it*-cleft can tell us about the nature of constructional change. As a result, this chapter contributes to the recent literature on diachronic construction grammar and grammaticalization theory. I concluded this chapter by showing how the historical evidence can be used to both inform and support my synchronic analysis of the *it*-cleft, resulting in a maximally explanatory account of the English *it*-cleft construction.

9.2 Final conclusions

This thesis offers a comparatively simple account of the structure and function of the English *it*-cleft. It argues in favour of a straightforward extraposition-from-NP analysis which views *it*-clefts in relation to other specificational copular constructions. However, the analysis is based upon an original semantic account of specificational meaning as involving an asymmetric predication relation which is dependent upon the inherent

semantics of definite noun phrases (rather than syntactic movement). By exploring and explaining the role that definite descriptions play in the creation of specificational meaning, the extraposition-from-NP (or discontinuous constituent) analysis of *it*-clefts gains considerable support and we are able to explain, rather than simply observe, the numerous similarities between *it*-clefts and definite noun phrases. Unlike most analyses of *it*-clefts, the account proposed in this thesis contains a historical component. This is in keeping with the fundamental objective of construction grammar that while motivation must be maximized, idiosyncrasies are nevertheless tolerated. Rather than explaining away the *it*-cleft's construction-specific attributes with a convoluted and overcomplicated synchronic account, this thesis makes use of historical evidence to explain how and why the *it*-cleft developed idiosyncratic properties, thereby informing and supporting an otherwise simple and intuitive analysis of the English *it*-cleft.

References

Adger, David and Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34, 325-359.

Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1, 149-168.

Atlas, Jay David and Stephen C. Levinson. 1981. *It*-clefts, informativeness, and logical form: Radical pragmatics (revised standard version). In Peter Cole (ed.) *Radical pragmatics*. New York: Academic Press, 1-62.

Ball, Catherine N. 1977. Th-clefts. *Pennsylvania Review of Linguistics* 2, 57-69.

Ball, Catherine N. 1991. *The historical development of the* it-*cleft*. Philadelphia PA: University of Pennsylvania PhD dissertation.

Ball, Catherine N. 1994a. The origins of the informative-presupposition *it*-cleft. *Journal of Pragmatics* 22, 603-628.

Ball, Catherine N. 1994b. Relative pronouns in *it*-clefts: The last seven centuries. *Language, Variation and Change* 6, 179-200.

Ball, Catherine N. 1996. A diachronic study of relative markers in spoken and written English. *Language, Variation and Change* 8, 227-258.

Ball, Catherine N. 1999. A short history of agreement in *it*-clefts. Paper presented at the Modern Language Association annual convention (session 434: New Work in Language Theory III), Chicago IL (29 December 1999).

Blom, Alied and Saskia Daalder. 1977. *Syntaktische theorie en taalbeschrijving*. Muiderberg: Coutinho.

Bolinger, Dwight. 1972. A look at equations and cleft sentences. In Evelyn S. Firchow, Kaaren Grimstad, Nils Hasselmo and Wayne A. O'Neil (eds.) *Studies for Einar Haugen; presented by friends and colleagues*. The Hague and Paris: Mouton, 96-114.

Bolinger, Dwight. 1977. *Meaning and form*. London and New York: Longman.

Borkin, Ann. 1984. *Problems in form and function*. Norwood NJ: Ablex Publishing Corporation.

Breivik, Leiv E. 1986. Some remarks on cleft sentences in present-day English. In Dieter Kastovsy and Aleksander Szwedek (eds.) *Linguistics across historical and geographical boundaries, Volume 2: Descriptive, contrastive and applied linguistics*. Berlin, New York and Amsterdam: Mouton de Gruyter, 815-826.

Brenier, Jason M. and Laura A. Michaelis. 2005. Optimization via syntactic amalgam: Syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1, 45-88.

Bybee, Joan. 1985. *Morphology: An inquiry into the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, Joan. 2003. Mechanisms of change in grammaticalization: The role of frequency. In Brian D. Joseph and Richard D. Janda (eds.) *The handbook of historical linguistics*. Oxford: Blackwell, 602-623.

Calude, Andreea S. 2007. *Demonstrative clefts in spoken English*. Auckland: University of Aukland PhD dissertation.

Calude, Andreea S. 2008. Demonstrative clefts and double cleft constructions in spontaneous spoken English. *Studia Linguistica* 62, 78-118.

Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow and Adrian Akmajian (eds.) *Formal Syntax*. New York: Academic Press.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge MA: MIT Press.

Clark, Herbert H. and Susan E. Haviland. 1977. Comprehension and the given-new contract. In Roy O. Freedle (ed.) *Discourse processes: Advances in research and theory, Volume 1: Discourse production and comprehension*. Norwood NJ: Ablex Publishing Corporation, 1-40.

Collins, Peter. 1991a. *Cleft and pseudo-cleft constructions in English*. London and New York: Routledge.

Collins, Peter. 1991b. Pseudocleft and cleft constructions: A thematic and informational interpretation. *Linguistics* 29, 481-519.

Croft, William. 1991. *Syntactic categories and grammatical relations*. Chicago IL: The University of Chicago Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford and New York: Oxford University Press.

Croft, William and D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Culicover, Peter W. and Michael S. Rochemont. 1990. Extraposition and the complement principle. *Linguistic Inquiry*. 21, 23-47.

Curme, George O. 1931. *Syntax*. New York: Heath.

Davidse, Kristin. 2000. A constructional approach to clefts. *Linguistics* 38, 1101-1131.

Declerck, Renaat. 1983. Predicational clefts. *Lingua* 61, 9-45.

Declerck, Renaat. 1984a. The pragmatics of *it*-clefts and WH-clefts. *Lingua* 64, 251-289.

Declerck, Renaat. 1984b. Some restrictions on clefts that highlight predicate nominals. *Journal of Linguistics* 20, 131-154.

Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press; Dordrecht: Foris.

Delahunty, Gerald P. 1982. *Topics in the syntax and semantics of English cleft sentences*. Bloomington IN: Indiana University Linguistics Club.

Delahunty, Gerald P. 1984. The analysis of English cleft sentences. *Linguistic Analysis* 13, 63-113.

Delin, Judy L. 1989. *Cleft constructions in discourse*. Edinburgh: University of Edinburgh PhD dissertation.

Delin, Judy L. 1992. Properties of *it*-cleft presupposition. *Journal of Semantics* 9, 289-306.

Den Dikken, Marcel, Andre Meinunger and Chris Wilder. 2000. Pseudoclefts and Ellipsis. *Studia Linguistica* 54, 41-89.

Den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge MA: MIT Press.

Den Dikken, Marcel. 2009. Predication and specification in the syntax of cleft sentences. Ms., CUNY Graduate Center.

De Swart, Henriette. 1998. Aspect shift and coercion. *Natural Language & Linguistic Theory* 16, 347-385.

Diewald, Gabriele. 2006. Context types in grammaticalization as constructions. *Constructions* SV1/9-2006, http://www.constructions-online.de (20 May 2009)

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75, 281-304.

É. Kiss, Katalin. 1998. Identification focus versus information focus. *Language* 74, 245-273.

Elffers, Els. 1979. De semantiek van de koppelwerkwoordzin en haar plaats in de taalbeschrijving. *Spektator* 9, 97-143.

Emonds, Joseph E. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.

Erdmann, Peter. 1986. A note on reverse *wh*-clefts in English. In Dieter Kastovsy and Aleksander Szwedek (eds.) *Linguistics across historical and geographical boundaries, Volume 2: Descriptive, contrastive and applied linguistics*. Berlin, New York and Amsterdam: Mouton de Gruyter, 851-858.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface.* Oxford: Oxford University Press.

Fillmore, Charles J., Paul Kay and Mary Kay O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64: 501-38.

Fillmore, Charles J., Paul Kay, Laura Michaelis and Ivan Sag. forthcoming. *Construction Grammar*. Stanford CA: CSLI Publications.

Fodor, Janet D. and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5, 355-398.

Fowler, Henry W. and Francis G. Fowler. 1908. *The King's English*, 2nd edn. Oxford: Clarendon Press; *Bartleby.com*. 1999. www.bartleby.com/116/ (17 November 2009)

Francis, Elaine J. 1999a. *Variation within lexical categories*. Chicago IL: University of Chicago PhD Dissertation.

Francis, Elaine J. 1999b. A conceptual semantic analysis of thematic structure in predicate nominals. Paper presented at the Texas Linguistics Society, Austin TX, http://web.ics.purdue.edu/~ejfranci/ejfrancis.htm (16 March 2009)

Francis, Elaine J. and Laura A. Michaelis. 2003. Mismatch: A crucible for linguistic theory. In Elaine J. Francis and Laura A. Michaelis (eds.) *Mismatch: Form-function incongruity and the architecture of grammar*. Stanford CA: CSLI publications, 1-27.

Francis, Elaine J. and Etsuyo Yuasa. 2008. A multi-modular approach to gradual change in grammaticalization. *Journal of Linguistics* 44, 45-86.

Fried, Mirjam. 2008. Constructions and constructs: Mapping a shift between predication and attribution. In Alexander Bergs and Gabriele Diewald (eds.) *Constructions and language change*. Berlin and New York: Mouton de Gruyter, 47-79.

Geluykens, Ronald. 1988. Five types of clefting in English discourse. *Linguistics* 26, 823-841.

Gisborne, Nikolas and Amanda L. Patten. forthcoming. Constructions and grammaticalization. In Bernd Heine and Heiko Narrog (eds.) *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago IL: Chicago University Press.

Goldberg, Adele E. 2003. Words by default: The Persian complex predicate construction. In Elaine J. Francis and Laura A. Michaelis (eds.) *Mismatch: Form-function incongruity and the architecture of grammar*. Stanford CA: CSLI publications, 117-146.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford and New York: Oxford University Press.

Goldberg, Adele E. and Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80, 532-568.

Gómez-González, María. 2004. A three-dimensional account of *it*-clefts in discourse: A corpus-based study. *Southwest Journal of Linguistics* 23, 81-120.

Grice, H. P. 1969. Vacuous names. In Donald Davidson and Jaakko Hintikka (eds.) *Words and objections*. Dordrecht: Reidel, 118-145.

Gundel, Jeanette K. 1977. Where do cleft sentences come from? *Language* 53, 543-559.

Gundel, Jeanette K. 1985. Shared knowledge and topicality. *Journal of Pragmatics* 9, 83-107.

Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 2001. Definite descriptions and cognitive status in English: why accommodation is unnecessary. *English Language and Linguistics* 5, 273-295.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English, part II. *Journal of Linguistics* 3, 199-244.

Halvorsen, Per-Kristian. 1978. The syntax and semantics of cleft constructions. *Texas Linguistics Forum 11*. Austin TX: Department of Linguistics, University of Texas.

Han, Chung-hye and Nancy Hedberg. 2008. Syntax and semantics of *it*-clefts: A Tree Adjoining Grammar analysis. *Journal of Semantics* 25, 345-380.

Hankamer, Jorge. 1974. On the non-cyclic nature of WH-clefting. *Chicago Linguistic Society* (*CLS*) 10, 221-233.

Hasselgård, Hilde. 2004. Adverbials in IT-cleft constructions. *Language & Computers* 49, 195-221.

Haugland, Kari E. 1993. A note on cleft and existential sentences in Old English. *English Studies* 5, 407-413.

Hawkins, John. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality predication*. Atlantic Highlands NJ: Humanities Press.

Hedberg, Nancy. 1990. *Discourse pragmatics and cleft sentences in English*. Minneapolis MN: University of Minnesota PhD dissertation.

Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76, 891-920.

Hedberg, Nancy and Lorna Fadden. 2007. The information structure of *it*-clefts, *wh*-clefts and reverse *wh*-clefts in English. In Nancy Hedberg and Ron Zacharski (eds.) *The grammar-pragmatics interface: Essays in honor of Jeanette K. Gundel*. Amsterdam and Philadelphia: John Benjamins, 49-76.

Heggie, Lorie A. 1988. *The syntax of copular structures*. Los Angeles CA: University of Southern California PhD dissertation.

Heim, Irene R. 1982. *The semantics of definite and indefinite noun phrases*. Amherst MA: University of Massachusetts PhD dissertation.

Heller, Daphna. 2002. On the relation of connectivity and specificational pseudoclefts. *Natural Language Semantics* 10, 243-284.

Heller, Daphna. 2005. *Identity and information: Semantic and pragmatic aspects of specificational sentences*. Newark NJ: The State University of New Jersey PhD dissertation.

Heycock, Caroline and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30, 365-397.

Heycock, Caroline and Antony Kroch. 2002. Topic, focus, and syntactic representations. In Line Mikkelsen and Chris Potts (eds.) *West Coast Conference on Formal Linguistics (WCCFL) 21*. Somerville MA: Cascadilla Press, 101-125.

Higgins, F. Roger. 1979. *The pseudo-cleft construction in English*. New York: Garland Publishing.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal? In Walter Bisang, Nikolaus Himmelmann and Bjorn Wiemer (eds.) *What makes grammaticalization? A look from its fringes and its components*. Berlin and New York: Mouton de Gruyter, 21-42.

Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.

Hopper, Paul J. 2001. Grammatical constructions and their discourse origins: prototype or family resemblance? In Martin Pütz, Suzanne Niemeier, and René Dirven (eds.) *Applied Cognitive Linguistics I: Theory and Language Acquisition* (Cognitive Linguistics Research, 19) Berlin and New York: Mouton de Gruyter, 101-129.

Huang, Yan. 2007. *Pragmatics*. Oxford and New York: Oxford University Press.

Huddleston, Rodney. 1984. *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Hudson, Richard A. 1990. *English Word Grammar*. Oxford: Blackwell.

Hudson, Richard A. 1997. The rise of auxiliary DO: Verb-non-raising or category strengthening? *Transactions of the Philological Society* 95: 41-72.

Hudson, Richard A. 2003. Mismatches in default inheritance. In Elaine J. Francis and Laura A. Michaelis (eds.) *Mismatch: Form-function incongruity and the architecture of grammar*. Stanford CA: CSLI publications, 269-317.

Hudson, Richard A. 2007. *Language networks: The new Word Grammar*. Oxford: Oxford University Press.

Israel, Michael. 1996. The way constructions grow. In Adele Goldberg (ed.) *Conceptual structure, discourse and language*. Stanford CA: CSLI, 217-230.

Jacobson, Pauline. 1994. Binding Connectivity in Copular Sentences. In Mandy Harvey and Lynn Santelmann (eds.) *Semantics and linguistic theory* (*SALT*) *IV*. Ithaca NY: Cornell University, 161-178.

Jackendoff, Ray. 1977. *X-bar Syntax: A study of phrase structure*. Cambridge MA: MIT Press.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge MA: MIT Press.

Jespersen, Otto. 1927. *A modern English grammar on historical principles, Part 3*. Heidelberg: Winter.

Jespersen, Otto. 1937. *Analytic syntax*. London: Allen and Unwin.

Jespersen, Otto. 1949. *A modern English grammar on historical principles, Part 7*. Copenhagan: Munksgaard.

Johansson, Mats. 2001. Clefts in contrast: A contrastive study of *it* clefts and *wh* clefts in English and Swedish texts and translations. *Linguistics* 39, 547-582.

Kay, Paul and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The *what's X doing Y?* construction. *Language* 75: 1-33.

Kearns, Kate. 2000. *Semantics*. Basingstoke: Macmillan.

Kiparsky, Paul and Carol Kiparsky. 1970. Fact. In Manfred Bierwisch and Karl E. Heidolph (eds.) *Progress in Linguistics*. The Hague: Mouton, 143-173.

Koops, Christian and Martin Hilpert. 2009. The co-evolution of syntactic and pragmatic complexity: diachronic and cross-linguistic aspects of pseudoclefts. In Talmy Givón and Masayoshi Shibatani (eds.) *Syntactic Complexity: Diachrony, acquisition, neuro-cognition, evolution*. Amsterdam: John Benjamins.

Kruisinga, Etsko. 1932. *A handbook of present-day English. Part 2: English accidence and syntax (vol 3)*, 5th edn. Groningen: Noordhoff.

Lahousse, Karen. 2009. Specificational sentences and the influence of information structure on (anti-)connectivity effects. *Journal of Linguistics* 45, 139-166.

Lakoff, George. 1987. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago IL: University of Chicago Press.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39, 463-516.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar, Vol I: Theoretical prerequisites*. Stanford CA: Stanford University Press.

Langacker, Ronald. 1990. Subjectification. *Cognitive Linguistics* 1: 5-38.

Langacker, Ronald. 1991. *Foundations of Cognitive Grammar, Vol II: Descriptive application*. Stanford CA: Stanford University Press.

Lehmann, Christian. 2008. Information structure and grammaticalization. In Elena Seoane and Maria Jose Lopez-Couso, (eds.) *Theoretical and empirical issues in grammaticalization.* Amsterdam and Philadelphia: John Benjamins, 207-229.

Lewis, David.1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8, 339-359.

Lightfoot, David W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David W. 1999. *The development of language: acquisition, change, and evolution*. Oxford: Blackwell.

Los, Bettelou. 2009. The consequences of the loss of verb-second in English: information structure and syntax in interaction. *English Language and Linguistics* 13, 97-125.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Matsunami, Tamotsu. 1961. A historical consideration of the disjunctive formula: *it is I that am to blame*. *Eibungaku Kenkyu* (*Studies in English literature*). Tokyo: English Literary Society of Japan. 38, 1-15.

Michaelis, Laura A. 2003. Headless constructions and coercion by construction. In Elaine J. Francis and Laura A. Michaelis (eds.) *Mismatch: Form-function incongruity and the architecture of grammar*. Stanford CA: CSLI Publications, 259-310.

Mikkelsen, Line. 2002. Two types of definite description subjects. In Malvina Nissim (ed.) *The European Summer School in Logic, Language and Information* (*ESSLLI*) *student session 7*, 1-13.

Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: John Benjamins.

Mikkelsen, Line. 2007. On so-called truncated clefts. In Ljudmila Geist and Björn Rothstein (eds.), *Kopulaverben under kopulasätze: Intersprachliche und intrasprachliche aspekte*. Tübingen: Niemeyer Verlag, 47-68.

Mitchell, Bruce. 1985. *Old English syntax, Volume 1: Concord, the parts of speech and the sentence*. Oxford: Clarendon Press.

Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.

Neale, Stephen. 1990. *Descriptions*. Cambridge MA: MIT Press.

Nelson, Gerald. 1997. Cleft constructions in spoken and written English. *Journal of English Linguistics* 25, 340-348.

Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14, 177-202.

Nunberg, Geoffrey, Thomas Wasow and Ivan Sag. 1994. Idioms. *Language* 70, 491-538.

O'Neil, Wayne. 1977. Clause adjunction in Old English. *General Linguistics* 17: 199-211.

Partee, Barbara H. 1986. Ambiguous pseudoclefts with unambiguous *be*. In Steve Berman, Jae-Woong Choe, and Joyce McDonough. *The North East Linguistic Society* (*NELS*) *16*. Amherst MA: University of Massachusetts, 354-366.

Patten, Amanda L. 2007. How specificational are cleft sentences? Paper presented at the Second International Conference of the Linguistics of Contemporary English (ICLCE2), University of Toulouse (3 July 2007) and the annual Linguistics Association of Great Britain meeting (LAGB), Kings College London (30 August 2007).

Patten, Amanda L. 2008. Accommodating predicational and proverbial clefts. Paper presented at the Cleft Workshop (Cleft08), Centre of General Linguistics, Berlin (29 November 2008).

Patten, Amanda L. in press. Grammaticalization and the *it*-cleft construction. In Elizabeth Closs Traugott and Graeme Trousdale (eds.) *Gradience, gradualness and grammaticalization*. Amsterdam: John Benjamins.

Pavey, Emma Louise. 2004. *The English* it*-cleft construction: A Role and Reference Grammar analysis*. Sussex: University of Sussex PhD dissertation.

Percus, Orin. 1997. Prying open the cleft. In Kiyomi Kusumoto (ed.) *The 27th Annual Meeting of the North East Linguistics Society*. Amherst MA: GLSA, 337-351.

Pinkham, Jessie and Jorge Hankamer. 1975. Deep and shallow clefts. *Chicago Linguistic Society* (*CLS*) *11*, 429-450.

Poutsma, Hendrick. 1928. *A grammar of late modern English. Part 1: The sentence. First half: The elements of the sentence*. Groningen: Noordhoff.

Prince, Ellen F. 1978. A comparison of *wh*-clefts and *it*-clefts in discourse. *Language* 54, 883-906

Radford, Andrew. 1997. *Syntactic theory and the structure of English*. Cambridge: Cambridge University Press.

Reeve, Matthew. 2008. English versus Russian: Two types of cleft construction. Paper presented at the Cleft Workshop (Cleft08), Centre of General Linguistics, Berlin (29 November 2008).

Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist approach to grammaticalization* (Cambridge Studies in Linguistics, 100). Cambridge: Cambridge University Press.

Rochemont, Michael. 1986. *Focus in generative grammar*. Philadelphia: John Benjamins.

Russell, Bertrand. 1905. On denoting. *Mind* 14, 479-93.

Rydén, Mats. 1966. *Relative constructions in early sixteenth century English*. Uppsala: Almqvist & Wiksell.

Schlenker, Phillipe. 2003. Clausal equations: A note on the connectivity problem. *Natural Language & Linguistic Theory* 21, 157-214.

Schmid, Hans-Jörg. 2001. 'Presupposition can be a bluff': How abstract nouns can be used as presupposition triggers. *Journal of pragmatics* 33, 1529-1552.

Sharvit, Yael. 1999. Connectivity in specificational sentences. *Natural Language Semantics* 7, 299-339.

Sharvit, Yael. 2003. Tense and identity in copular constructions. *Natural Language Semantics* 11, 363-393.

Sornicola, Rosanna. 1988. It-clefts and wh-clefts: Two awkward sentence types. *Journal of Linguistics* 24, 343-379.

Stanley, Jason and Zoltán Gendler Szabó. 2000. On quantifier domain restriction. *Mind & Language* 15, 219-261.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam and Philadelphia: John Benjamins.

Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Winifred P. Lehmann and Yakov Malkiel (eds.) *Perspectives on historical linguistics*. Amsterdam: Benjamins, 245-271.

Traugott, Elizabeth Closs. 1988. Pragmatic Strengthening and Grammaticalization. In Shelley Axmaker, Annie Jaisser, and Helen Singmaster (eds.) *The Berkeley Linguistics Society* (*BLS*) *14*, 406-416.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 57, 33-65

Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In Brian D. Joseph and Richard D. Janda (eds.) *The handbook of historical linguistics*. Oxford: Blackwell, 624-647.

Traugott, Elizabeth Closs. 2007. The concepts of constructions mismatch and type-shifting from the perspective of grammaticalization. *Cognitive Linguistics* 18: 513-557.

Traugott, Elizabeth Closs. 2008. "All that he endeavoured to prove was…": On the emergence of grammatical constructions in dialogic contexts. In Robin Cooper and Ruth Kempson (eds.) *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. London: Kings College Publications, 143-177.

Trousdale, Graeme. 2008a. A constructional approach to lexicalization processes in the history of English: Evidence from possessive constructions. *Word structure* 1: 156-177.

Trousdale, Graeme. 2008b. Constructions in grammaticalization and lexicalization: Evidence from the history of a composite predicate construction in English. In Graeme Trousdale and Nikolas Gisborne (eds.) *Constructional approaches to English Grammar*. Berlin: Mouton de Gruyter, 33-67.

Trousdale, Graeme. 2008c. Grammaticalization, constructions and the grammaticalization of constructions. Paper presented at New Reflections in Grammaticalization 4, KU Leuven (16 July 2008). Ms., University of Edinburgh.

Vallduví, Enric and Elisabet Engdahl. 1996. The linguistic realization of information packaging. *Linguistics* 34: 459-519.

Visser, F. Th. 1970. *An historical syntax of the English language. Part 1: Syntactical units with one verb*. Leiden: Brill.

Ward, Gregory, Betty Birner and Rodney Huddleston. 2002. Information packaging. In Huddleston, Rodney and Geoffrey Pullum (eds.) *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press, 1363-1447.

Warner, Anthony. 1993. *English Auxiliaries: Structure and history*. Cambridge: Cambridge University Press.

Watson, Nicholas and Jacqueline Jenkins (eds.). 2006. *The writings of Julian of Norwich: A vision showed to a devout woman and a revelation of love*. Pennsylvania: The Pennsylvania State University Press.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11, 203-238.

Williams, Edwin. 1983. Semantic vs. syntactic categories. *Linguistics and Philosophy* 6: 423-446.

Williams, Edwin. 1994. *Thematic structure in syntax*. Cambridge MA: MIT Press.

Wirth, Jessica R. 1978. The derivation of cleft sentences in English. *Glossa* 12, 58-81.